

# **Acta Psychologica**

EUROPEAN JOURNAL OF PSYCHOLOGY

Vol. 9 - 1953



SWETS & ZEITLINGER B.V.

AMSTERDAM - 1975





# Acta Psychologica

INCLUDING

NETHERLANDS-SCANDINAVIAN, BELGIUM AND  
SWITZERLAND JOURNAL OF PSYCHOLOGY

EDITOR

G. RÉVÉSZ

AMSTERDAM

CO-EDITORS

Sir FREDERIC BARTLETT, Cambridge  
H. J. F. W. BRUGMANS, Groningen  
E. KAILA, Helsinki - † D. KATZ,  
Stockholm - A. E. MICHOTTE, Leuven  
J. PIAGET, Paris-Genève - † E. RUBIN,  
Kjöbenhavn - H. SCHJELDERUP, Oslo



Volume IX

1953

NORTH-HOLLAND PUBLISHING COMPANY  
AMSTERDAM

# Acta Psychologica

EDITED BY

W. J. VAN DER KOOIJ, D. SCHEFFÉ, and  
J. VAN DER LINDEN, D. SCHEFFÉ

Volume 1

1972

AMSTERDAM

1972

ACTA PSYCHOLOGICA  
A JOURNAL OF  
PSYCHOLOGY  
AND  
PSYCHIATRY  
EDITED BY  
W. J. VAN DER KOOIJ, D. SCHEFFÉ,  
AND J. VAN DER LINDEN, D. SCHEFFÉ

1972

ACTA PSYCHOLOGICA

1972

ACTA PSYCHOLOGICA

1972

B  
B  
C  
C  
F  
H  
H  
K  
K  
L  
N  
O  
P  
R  
—  
—  
R  
S  
V  
Z

## CONTENTS

	Page
BEVAN JR, WILLIAM and GRETHA SEELAND, An Exploration of the Influence of Personal Relevance upon Statements of Aesthetic Preference . . . . .	274
BLUMENFELD, WALTER, The Precision of the "Black Thread Method" and Weber's Law . . . . .	201
CATTELL, RAYMOND B., A Quantitative Analysis of the Changes in the Culture Pattern of Great Britain 1837-1937, by P-Technique . . . . .	99
COHEN, JOHN, Social Thinking . . . . .	146
FRIJDA, NICO H., The Understanding of Facial Expression of Emotion . . . . .	294
HUMPHREY, G., Some Reflections upon Gilbert Ryle's Considerations	197
HUSÉN, TORSTEN, The Stability of Intelligence Test Scores . . . .	53
KAISER, L., Contribution to the Psychologic and Linguistic Value of Melody . . . . .	288
KATZ, DAVID, Psychologie des Sicherheitsmarginals . . . . .	255
LENNEP, D. J. VAN and R. H. HOUWINK, Projection Tests and Overt Behavior . . . . .	240
NYSSSEN, R. et J. HOZAY, Essais d'explication de l'erreur caractéristique dans un mode de délimitation des régions cutanées	122
ORMIAN, HAIM, Child Psychology and "Controversy of Schools" . .	16
PUMPHREY, R. J., The Origin of Language . . . . .	219
RÉVÉSZ, G., David Katz . . . . .	97
———, In Memoriam Gustav Kafka . . . . .	183
———, Edgar Rubin . . . . .	254
RYLE, GILBERT, Thinking . . . . .	189
SANDSTRÖM, CARL IVAR, Sex Differences in Localization and Orientation . . . . .	82
VALPOLA, VELI, Über das Messen in der Psychologie . . . . .	1
ZEEMAN, W. P. C. and C. OTTO ROELOFS, Some Aspects of Apparent Motion . . . . .	159



# ÜBER DAS MESSEN IN DER PSYCHOLOGIE

VON

VELI VALPOLA (Helsinki)

## I

Es lässt sich nicht leugnen, dass die Biologie und die Psychologie als Wissenschaften auf einem bedeutend niedrigeren Niveau stehen als z.B. die Physik und die sog. exakten Naturwissenschaften im allgemeinen. Die letztgenannten haben eine Technik erschaffen, die die Relationen der lebendigen Dinge zueinander und zur leblosen Natur völlig umgestürzt hat, während die biologische und psychologische Technik sich in ihren ersten Anfangsstadien befindet. Auch in theoretischer Hinsicht ist der Unterschied sehr deutlich. Die Ergebnisse der exakten Naturwissenschaften erscheinen in der Form von Theorien, deren logische Struktur ausserordentlich kompliziert und reich ist, während die Ergebnisse der Biologie und Psychologie in einer trivialen und logisch einfachen Sprache ausgedrückt werden, die die Alltagssprache kaum überschreitet.

Diese ausgeprägte, sowohl praktische als theoretische Verschiedenheit beruht vor allem darauf, dass die Mathematik innerhalb der exakten Naturwissenschaften eine gänzlich verschiedene Rolle als innerhalb der anderen Naturwissenschaften spielt, was schon bei einer oberflächlichen Betrachtung in die Augen fällt. Bei der Anwendung der Mathematik ist die Verwendung von ganzen Zahlen nicht von grosser Bedeutung, denn eigentlich kann alles, worüber man nur sprechen kann, auch abgezählt werden. Es sind erst die *Messungen*, die die Mathematik zu einem wirklich leistungsfähigen Mittel machen, zu einem Mittel wie es z.B. in der mathematischen Physik in die Erscheinung tritt. Die Entdeckung einer messbaren Quantität eröffnet die Möglichkeit, die reellen Zahlen in der Form von Masszahlen mit der Wirklichkeit zu verknüpfen (die Masszahlen sind wesentlich reelle Zahlen, unter denen auch rationale und ganze Zahlen vorkommen). Erst dadurch können die Begriffe der Theorie der stetigen Funktionen und der Infinitesimalrechnung benutzt werden, und wir wissen alle in welcher

geradezu überraschenden Weise diese Begriffe mit der Wirklichkeit harmonisieren. Dadurch können in günstigen Fällen unüberschaubare Mannigfaltigkeiten in Formeln von etwa nur zehn Zeichen kondensiert werden.

Man ist schon lange darüber im klaren gewesen, dass gerade die Messungen und die darauf beruhenden Anwendungen der Mathematik die exakten Naturwissenschaften so hoch über die anderen Naturwissenschaften erhoben haben. Es ist also kein Wunder, dass man auch innerhalb der Psychologie — seit Weber und Fechner — eifrig versucht hat, Messungen auszuführen und bei den psychologischen Forschungen zu benutzen. Die bisherigen Resultate sind jedoch gewissen Bedenken ausgesetzt, und es herrschen Meinungsverschiedenheiten über die Methodik und die Berechtigung der psychologischen Messungen. Misstrauen erregt besonders der Sachverhalt, dass man in der Psychologie oft augenfällig der physikalischen Methodik nachahmen will. Es ist auch auffallend, dass gewisse sehr komplizierte und umständliche statistische Methoden bei den psychologischen Messungen in der letzten Zeit grosse Verbreitung gefunden haben; ich meine hier in erster Linie die sog. „Faktorenanalyse“. Bei den grundlegenden physikalischen Messungen dagegen ist nichts derartiges vonnöten, sondern alles ist ganz einfach (nämlich in theoretischer Hinsicht, praktische Schwierigkeiten kommen selbstverständlich vor). Trotz des eifrigen und unermüdlichen Rechnens sind die mathematischen Ergebnisse der Psychologie sehr gering. Die Sachlage bedarf offensichtlich einer Klarlegung.

Es ist nun meine Absicht, das Problem des Messens von logischem und philosophischem Standpunkt aus zu beleuchten zu versuchen und die Stellung der Messungen im Gebiet der empirischen Begriffe zu behandeln <sup>1)</sup>.

## II

Bei der Messung wird eine grosse Menge von Objekten miteinander verglichen, um zu ermitteln, wie „viel“ oder in wie

---

<sup>1)</sup> Gedankengänge, die nahe verwandt mit denjenigen dieses Aufsatzes sind, sind auch von Kai von Fieandt („*Kan psykometriken ersätta holitisk personlighetsdiagnos*“, Nordisk psykologi 1950 S. 41—47 mit englischer Zusammenfassung) und Pentti Ikonen („*Some Methodological Problems in Test Psychology*“, *Acta Psychologica Fennica* I, 1951, S. 30—46), vorgelegt worden.



„hohem Grade“ sie die jeweils in Frage stehende Quantität besitzen. Es handelt sich erstens um die Bestimmung der Grössenordnung der betreffenden Objekte und zweitens um die Darstellung der quantitativen Unterschiede mit Hilfe von Zahlen. Diese Operationen zusammen bilden die Messung.

Die quantitative Vergleichung beruht auf gewissen Relationen, die „gegeben“ sein müssen, damit man überhaupt von einer Messung und einer Quantität sprechen könne. Zu jeder Quantität gehören verschiedene, gerade für sie charakteristische Relationen, die bei verschiedenen Quantitäten im allgemeinen voneinander getrennte Relationen sind. Diese für die Quantität grundlegenden Relationen können bei jeder herkömmlich anerkannten empirisch messbaren Quantität in der Erfahrung gefunden werden.

Die Relationen, um die es sich bei einer Quantität handelt, verknüpfen die betrachteten Objekte miteinander. Diese Objekte bilden einen Bereich oder eine Klasse, derart, dass jedes Element der Klasse mit jedem anderen Element derselben Klasse quantitativ vergleichbar ist. Zu jeder Quantität gehört also eine vollständig bestimmte Klasse, und wir können sagen, dass jede Quantität eine entsprechende Eigenschaft (Qualität) bestimmt, und zwar diejenige Eigenschaft, die die und nur die Objekte besitzen, die bezüglich jener Quantität miteinander vergleichbar sind; die Klasse dieser Objekte ist mit jener Eigenschaft logisch gleichwertig. — Einige Beispiele: Zu der Quantität, die man das Gewicht nennt, gehört die Klasse der materiellen Körper, denn diese und nur diese besitzen Gewicht; zu der Quantität Zeit gehört die Klasse der Ereignisse (als dimensionslose „Zeitpunkte“ aufgefasst); zu der Quantität Länge gehört die Klasse der starren Körper, denn starre und nur starre Körper haben eine bestimmte räumliche Länge; zu den gewöhnlichen psychologischen Quantitäten wie z.B. dem sprachlichen Auffassungsvermögen oder dem räumlichen Vorstellungsvermögen gehört die Klasse der Menschen.

Als erste unter den grundlegenden quantitativen Relationen können wir die sog. *Gleichheit* ansehen; diese Relation besteht zwischen zwei Elementen der Klasse, falls und nur falls sie hinsichtlich der betreffenden Quantität gleich sind. Zu jeder Quantität gehört eine bestimmte Gleichheitsrelation. — Beispiele: Beim Gewicht sagen wir, dass zwei Körper gleich

schwer sind; bei Zeit, dass zwei Ereignisse gleichzeitig sind; bei der Intelligenz, dass zwei Menschen gleich intelligent sind. Für jede Quantität gibt es (oder sollte es geben) ein eindeutiges empirisches Kriterium, wodurch man in jedem gegebenen Falle über die quantitative Gleichheit bzw. Ungleichheit zweier Elemente der betreffenden Klasse entscheiden kann.

Die zweite Relation ist die *ordnende Relation*, die zwischen zwei Elementen  $x$  und  $y$  der Klasse besteht, falls und nur falls  $x$  quantitativ kleiner als  $y$  ist (man sagt auch:  $x$  besitzt die betreffende Quantität weniger oder mit schwächerer Intensität als  $y$ ). — Beispiele: Ein Körper ist leichter als ein anderer, ein Ereignis ist früher als ein anderes, ein Mensch ist weniger intelligent als ein anderer.

Diese zwei Relationen bestimmen eine quantitative Ordnung der Elemente, die „Grössenordnung“. Diese ist eine reihenartige Anordnung, wie sie in der Anordnung der Punkte auf der geometrischen Geraden anschaulich zum Vorschein kommt; sie ist nicht zyklisch und nicht verzweigend, sondern sie muss eine sog. eindimensionale Anordnung sein. In solcher reihenartigen Anordnung gehört jedes Element der betreffenden Klasse zu einer eindeutig determinierten Stelle; zu einer und derselben Stelle der Anordnung gehören die und nur die Elemente, die quantitativ gleich sind, und ein Element gehört zu einer früheren Stelle als ein anderes, falls und nur falls es quantitativ kleiner ist als dieses.

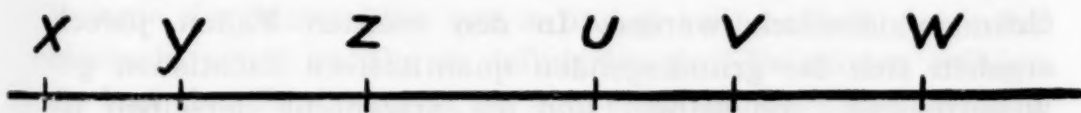
Für die Messbarkeit der Elemente ist aber eine quantitative Ordnung nicht hinreichend. Die Grössenverhältnisse der Elemente können zwar durch Zahlen dargestellt werden, aber auf unendlich viele wesentlich verschiedene Weisen. Vergleichen wir z.B. die Zahlenreihen 1, 2, 3, 4, 5, . . . . . 1, 4, 9, 16, 25, . . . . und 1, 2, 4, 5, 9, 10, 16, 17, . . . . , so sehen wir, dass die Grössenordnung in allen Reihen dieselbe ist; die quantitativen Strukturen sind aber verschieden.

Damit man von einer messbaren Quantität und einer messbaren Struktur mit voller Berechtigung sprechen könnte, ist eine „Grössenordnung“, eine quantitative reihenartige Ordnung, nicht hinreichend. Es muss in der Erfahrung notwendig eine weitere Relation gegeben sein, die das System der Masszahlen so weit bestimmt, dass durch die Festsetzung der Masseinheit und des Nullpunktes die Masszahlen eindeutig bestimmt werden. Gewöhn-



lich wird man hier die Relation der Summe (Addition) hervorheben, d.h. die Relation, die zwischen drei Elementen  $x$ ,  $y$  und  $z$  besteht, falls und nur falls die Summe von  $x$  und  $y$  quantitativ gleich mit  $z$  ist. Auch einige andere Relationen wären hier hinreichend, u.a. diejenige, die zwischen zwei Elementen besteht, falls und nur falls das eine das Zweifache des anderen ist, sowie auch diejenige, die zwischen drei Elementen besteht, falls und nur falls eines von diesen gerade in der Mitte zwischen den zwei anderen liegt. Die logisch einfachste von diesen für die Messbarkeit hinreichenden Relationen, die (annähernd) dieselbe Rolle spielt wie die oben erwähnten, ist die sog. *Distanzgleichheit*. Diese Relation besteht zwischen vier Elementen  $x$ ,  $y$ ,  $z$  und  $u$  der Klasse, falls und nur falls die quantitative Distanz von  $x$  bis  $y$  gleich der quantitativen Distanz von  $z$  bis  $u$  ist. Die Distanzgleichheit ist der geometrischen Streckenkongruenz teilweise ähnlich; es ist jedoch vorteilhaft, die Distanzen als „gerichtete Grössen“, als „Vektoren“ zu behandeln (so dass die Distanz von  $x$  bis  $y$  von der Distanz von  $y$  bis  $x$  verschieden ist, wenn  $x$  und  $y$  nur quantitativ ungleich sind).

Die für die Messung grundlegenden drei Relationen sind nicht voneinander logisch unabhängig, sondern sie müssen ganz bestimmten logischen Forderungen genügen, um die Rede von einer messbaren Quantität berechtigt zu machen. Die Relationen müssen in der Klasse der Elemente eine bestimmte logische Struktur erzeugen, d.h. es muss zwischen ihnen eine bestimmte logische Verbindung bestehen. Als Beispiele von diesen logischen Forderungen seien die folgenden zwei erwähnt: die Distanzen von  $x$  bis  $y$  und von  $x$  bis  $z$  sind quantitativ gleich, falls und nur falls  $y$  und  $z$  quantitativ gleich sind und in der Ordnung zu gleicher Stelle gehören (es ist zu bemerken, dass die Gleichheit der Elemente eine andere Relation ist als die Gleichheit der Distanzen); falls die Distanzen von  $x$  bis  $y$  und von  $u$  bis  $v$  miteinander gleich sowie die Distanzen von  $y$  bis  $z$  und von  $v$  bis  $w$  miteinander gleich sind, so sind die Distanzen von  $x$  bis  $z$  und von  $u$  bis  $w$  miteinander gleich.



Diese logischen Forderungen können hier nicht näher behan-

delt werden, sondern ich muss mich damit begnügen, auf meinen Aufsatz „Über den Begriff der Quantität“ (in *Ajatus*, dem Jahrbuch des finnischen philosophischen Vereins, 15 (1948), Seiten 261-298) hinzuweisen, wo die erwähnten logischen Forderungen exakt und vollständig dargestellt sind. — Das anschaulichste Beispiel von einer messbaren Struktur ist die Anordnung der Punkte auf einer geometrischen Geraden, wobei die Kongruenz- und Stetigkeitsaxiome mit berücksichtigt werden müssen.

Bei einem einzigen Objekt kann unmöglich von einer Quantität gesprochen werden. Der Begriff der Quantität setzt eine umfangreiche Klasse von Elementen voraus, die miteinander vergleichbar sind. Das Wesentlichste bei der Quantität sind nicht die einzelnen „Größen“ oder ihre Masszahlen und auch nicht die empirischen Relationen, auf Grund derer die Masszahlen erhalten werden. Das Wesentlichste ist die logische Verbindung, die zwischen den grundlegenden quantitativen Relationen besteht. Wir können also sagen, dass *die Quantität wesentlich eine Relationsstruktur von ganz bestimmter Art ist, welche Relationsstruktur in der Klasse der Elemente durch die grundlegenden Relationen gleichsam erzeugt wird.*

Zwei messbaren Quantitäten, die zu einer gewissen Klasse von Elementen gehören, sind eine und dieselbe, falls und nur falls die entsprechenden Relationsstrukturen identisch sind, und zwar logisch, notwendig, identisch; falls die Nullpunkte und die Mass-einheiten dann identisch sind, so werden auch die Masszahlen identisch sein. Hier genügt also keine empirisch festgestellte Gleichheit, sondern sie muss aus den Definitionen der grundlegenden quantitativen Relationen folgen und unabhängig von der Erfahrung bestehen. Eine bloss empirisch festgestellte Gleichheit zweier quantitativer Strukturen bedeutet nicht die Identität der entsprechenden Quantitäten, sondern lediglich eine empirische Regelmässigkeit, die jedoch sehr wertvoll sein kann.

Es ist zu bemerken, dass eine und dieselbe Quantität auch durch verschiedene Relationen bestimmt werden kann, falls nämlich die entsprechenden Relationsstrukturen aus logischen Gründen identisch werden. In den meisten Fällen jedoch ergeben sich die grundlegenden quantitativen Relationen gewissermassen „von selbst“, und die Erwählung derselben ist ziemlich frei von Willkür.

### III

Wenn die erforderlichen quantitativen Relationen zur Verfügung stehen und wenn empirisch festgestellt worden ist, dass sie den gehörigen logischen Forderungen genügen, dann lässt es sich beweisen, dass jedes Element der Klasse gemessen werden kann. Werden der „Nullpunkt“ und die „Masseinheit“ festgesetzt, so wird die Masszahl jedes Elementes eindeutig determiniert und kann auch durch empirische Operationen mit beliebiger Genauigkeit bestimmt werden.

Die Messung ergibt für jedes Element der Klasse eine eindeutig bestimmte Masszahl. Dieses ist der logische Kern des Messens, nebensächlich ist dabei die Methode, wodurch die Masszahlen erhalten worden sind. Gewöhnlich spricht man allerdings — wie auch oben der Fall war — von gewissen empirischen Relationen, auf deren Anwendung die Messung sich gründet und die die Quantität bestimmen. Von fundamentaler Wichtigkeit aber ist es einzusehen, dass auch die Masszahlen die Quantität vollständig bestimmen. Es ist durchaus gleichgültig, auf welchem Wege die Masszahlen erhalten worden sind — sie bestimmen in jedem Falle die Relationsstruktur ebenso vollständig wie die quantitativen Relationen. Diese Relationsstruktur, die im System der Masszahlen zum Ausdruck kommt, ist das Wesentlichste bei der Quantität, und sie ist jedenfalls das Einzige, was bei der Anwendung der Mathematik notwendig ist.

Auf Grund der Messung erhält man die Masszahlen, auf Grund der Masszahlen aber können umgekehrt solche Relationen definiert werden, die die zugehörigen logischen Forderungen erfüllen. Es genügt, die folgenden drei Übereinkünfte zu treffen: erstens, zwei Elemente der betreffenden Klasse seien quantitativ gleich, falls und nur falls ihre Masszahlen arithmetisch gleich sind; zweitens, ein beliebiges Element  $x$  der Klasse sei quantitativ kleiner als ein Element  $y$ , falls und nur falls die Masszahl von  $x$  arithmetisch kleiner als die Masszahl von  $y$  ist; drittens, die quantitative Summe von zwei Elementen  $x$  und  $y$  sei quantitativ gleich einem Element  $z$ , falls und nur falls die arithmetische Summe der Masszahlen von  $x$  und  $y$  arithmetisch gleich der Masszahl von  $z$  ist. Die so definierten drei Relationen erfüllen die logischen Forderungen, die man einer quantitativen Struktur stellen kann. Vermittels der erhaltenen Relationen können wir



dann, wenn wir es wünschen, die Messung der Elemente der Klasse auf gewöhnlichem Wege durchführen. Werden der Nullpunkt und die Masseinheit passend gewählt, so ergibt die Messung gerade diejenigen Zahlen als Masszahlen, die die Grundlage für die Konstruktion der quantitativen Relationen bildeten.

Wir stellen fest, dass man für die Messung keine anderen logischen Forderungen stellen kann, als dass jedes Element eine eindeutig determinierte Masszahl erhalten soll. Die gehörige logische Struktur einer messbaren Quantität und die gehörigen quantitativen Relationen werden sich dann von selbst ergeben. Dieses führt zu dem Schluss, dass *die Elemente jeder (hinreichend grossen) Klasse gemessen werden können*, denn Masszahlen können wir für sie jedenfalls erhalten, auf einem oder anderem Wege.

Masszahlen können wir den Elementen z.B. durch eine Auslosung zuordnen. Man legt die Lotteriescheine der Elemente in einen Behälter und die der Zahlen in einen anderen und zieht dann wiederholt aus beiden Behältern gleichzeitig einen Lotterieschein aus. Dadurch erhält man für jedes Element eine eindeutig bestimmte Zahl, die man für die Masszahl erklärt. Bei einer Auslosung kann man allerdings nur endliche Mengen behandeln, dasselbe gilt aber auch für die gewöhnliche Messung. Bei der Auslosung erhält jedes beliebige Element eine eindeutig bestimmte Masszahl, indem man nötigenfalls neue Lotteriescheine in die Behälter legt. — Man könnte hier vielleicht einwenden, dass man bei wiederholter Auslosung ja ganz verschiedene Masszahlen erhalten würde. Dazu sei bemerkt, dass man die Wiederholbarkeit der Messung nicht allgemein fordern kann. In der Physik kommen zahlreiche Messungen vor, die nicht wiederholt werden können; z.B. die Messung der Temperaturen kann nicht wiederholt werden, da die Temperaturen der Körper sich mit der Zeit ändern; astronomische Messungen können nicht wiederholt werden, da die Himmelskörper sich bewegen; sogar die Zeit selbst wird gemessen, obwohl sie sich stetig ändert. Es sei noch die Bemerkung gemacht, dass man ein Protokoll über die Auslosung aufnehmen und dadurch eine invariante Grundlage für die „Messung“ erhalten könnte. Auch könnte man an jedem gemessenen Element einen Zettel anbinden, der dann die invariante Masszahl angeben würde.

Es muss hier noch ein Punkt behandelt werden, der oft Missverständnisse verursacht hat. Ich habe oben durchweg von der Messung von Elementen einer Klasse gesprochen. Im gewöhnlichen Sprachgebrauch, sowohl im alltäglichen als im wissenschaftlichen, wird jedoch von der Messung der Eigenschaften gesprochen (z.B. von der Messung des Gewichtes, der Temperatur, der Intelligenz u.s.w.). Wie verhält es sich damit?

Wir haben festgestellt, dass die hinreichende und notwendige Bedingung für die Messbarkeit die Existenz einiger, bestimmten logischen Forderungen genügenden Relationen ist. Mit derjenigen Eigenschaft, die man vermittels dieser Relationen zu messen glaubt, stehen die Relationen in keiner anderen Verbindung als dass jedes Element der betreffenden Klasse jene Eigenschaft besitzen muss. Es ist selbstverständlich, dass dieses nicht hinreichend ist, den Sinn des Ausdrucks „Messung einer Eigenschaft“ festzusetzen; eine und dieselbe Eigenschaft könnte ja auf viele verschiedene Weisen „gemessen“ werden, die voneinander gänzlich unabhängig sind.

Nehmen wir als Beispiel die sog. „Messung des Gewichtes“, die vermittels der Wage vor sich geht. Kann jemand angeben, weshalb die durch Benutzung der Wage erhaltenen empirischen Relationen gerade das Gewicht messen und nicht z.B. die Temperatur? Selbstverständlich wird mit der Wage das Gewicht gemessen, das will ich nicht bestreiten. Ich will nur klarlegen, dass der Zusammenhang der „zu messenden“ Eigenschaft und der Operation des Messens auf blosser Intuition beruht, eine rein erlebnismässige Selbstverständlichkeit ist, die ohne jeden sachlichen Grund dasteht.

Die Rede von der Messung der Eigenschaften ist ohne jeden präzisen Sinn. Ist eine Eigenschaft ein wirklich präziser und klarer Begriff, so besitzt jedes beliebige Objekt diese Eigenschaft oder besitzt sie nicht. Hinsichtlich derjenigen Elemente, die diese Eigenschaft besitzen, kann man nicht von einer kleineren oder grösseren „Menge“ (oder „Grad“ oder „Intensität“) sprechen.

In der psychologischen Literatur begegnet man oft der Auffassung, dass die Frage von der richtigen oder falschen Messung der psychologischen Eigenschaften ein fundamentales psychologisches Problem sei. Wir stellen fest, dass dieses Problem unlösbar ist, da die Frage von der richtigen oder falschen Messung einer Eigenschaft sinnlos ist. Wesentlich und notwendig

bei dem Begriffe der Quantität ist eine Relationsstruktur bestimmter Art, welche Struktur rein logisch ist und an und für sich keiner empirischen Charakterisierungen oder Bestimmungen fähig ist. Wollte man irgend ein Kriterium entdecken, wodurch man entscheiden könnte, ob ein bestimmtes Testsystem z.B. die Intelligenz oder etwa das räumliche Vorstellungsvermögen misst, so muss man sagen, dass es aus logischen Gründen kein solches Kriterium gibt.

Die häufig diskutierte Frage, welches Testsystem das richtige Mass der Intelligenz sei, ist sinnlos. Die Entscheidung zwischen verschiedenen Testsystemen kann nur auf einer unsicheren und ganz subjektiven Intuition beruhen. Zwei verschiedene Testsysteme bestimmen zwei verschiedene Quantitäten, so dass man sagen kann, sie „messen“ verschiedene Sachen. Es kann wohl vorkommen, dass zwei Testsysteme zu übereinstimmenden Ergebnissen führen, dies bedeutet aber nicht, dass sie eine und dieselbe Quantität messen würden, sondern es handelt sich lediglich um eine empirische Regelmässigkeit.

#### IV

Wir sind zu dem Schluss gelangt, dass man nur von der Messung der Elemente in einer Klasse (oder von der Messung der Klassen) sprechen kann. Wir haben auch festgestellt, dass die Quantität ein rein logischer Begriff ist (obwohl die Einzelheiten nicht näher behandelt werden konnten). Über die Messbarkeit und die Messung lässt sich nichts speziell Empirisches sagen, nichts anderes, als was die logischen Bestimmungen enthalten. Daraus folgt, wie schon erwähnt, dass jede hinreichend grosse Klasse messbar ist, sogar auf unendlich viele, wesentlich verschiedene Weisen.

Die Messungen und die messbaren Quantitäten sind von mannigfachster Art, und das Einzige, was sie gemeinsam haben, ist eine und dieselbe logische Struktur. Es besteht kein empirisches Kriterium, wodurch man die messbaren Quantitäten und besonders die zugehörigen quantitativen Relationen charakterisieren könnte. Bei den meisten physikalischen Quantitäten ist z.B. die quantitative Gleichheit bzw. Ungleichheit zweier Elemente allerdings eine ziemlich einfache und schlichte Sache, die sich sozusagen „von selbst herausstellt“. Der Versuch, eine

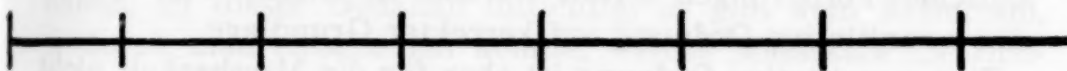


Begründung für diese intuitive Selbstverständlichkeit anzugeben, wird sich aber als vollständig unfruchtbar zeigen.

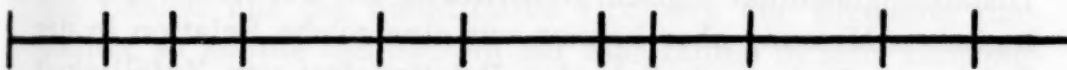
Da der Begriff der Quantität ein logischer Begriff ist, wird er ein empirischer Begriff erst dann, wenn den grundlegenden Relationen ein empirischer Gehalt zugelegt worden ist. Wie bei den anderen logischen Begriffen, liegt auch bei dem Begriffe der Quantität sein Wert und Belang ausschliesslich darin, worauf er angewendet wird. Es ist klar, dass die Masszahlen, falls sie durch eine Auslosung oder irgend ein ebenso zufälliges Verfahren erhalten werden, gänzlich wertlos und nutzlos sind. Sind die für die Messung grundlegenden Relationen für die Realität wichtig, so ist auch die Messung wertvoll, sind sie wertlos oder künstlich, so ist auch die Messung wertlos. Hier liegt der philosophische Kern im Problem des Messens.

Bei den grundlegenden physikalischen Messungen stehen jene Relationen in Verbindung mit zentralen Invarianzen der Erfahrung. Die Messung des Raumes ist die wichtigste Messung gerade darum, weil sie sich auf den innersten Kern unseres Realitätsbegriffs, den Begriff des starren Körpers, gründet.

Betrachten wir ein konkretes Beispiel. Führen wir eine Raummessung mittels eines bestimmten Massstabes aus: Wir erhalten dadurch eine Reihe von Punkten, die in *regelmässigen* Abständen voneinander liegen:



Vergleichen wir damit die untenstehende Figur, wo die Punkte *unregelmässig* und *zufällig* liegen:



Die erste Figur stellt eine „richtige“ Raummessung dar, die zweite eine „falsche“. Die logische Struktur hingegen ist bei beiden Messungen eine und dieselbe; in beiden Fällen wird jeder Punkt (wenn wir die Operation als vollführt denken) eine eindeutig bestimmte Masszahl erhalten. Welches ist denn der Unterschied zwischen diesen beiden Fällen — dieser Unterschied scheint ja ganz wesentlich und ausgeprägt zu sein? Es ist klar, dass im ersten Falle die Messung von dem benutzten bestimmten Massstab unabhängig ist, im zweiten Falle dagegen von einer

einmaligen und zufälligen Verteilung der Punkte abhängt. Im ersten Falle besteht folglich die empirische Gesetzmässigkeit, dass *jeder beliebige* starre Massstab zu *einer und derselben* Verteilung der Punkte führt, im zweiten Falle dagegen ist die Verteilung durch eine spezielle Figur bestimmt; es steht also eine grosse Allgemeinheit einem einzelnen Spezialfall gegenüber.

## V

Wir kommen nun zur anfangs gestellten Frage von den psychologischen Messungen. Hier ist die Sachlage nicht ganz klar. In den meisten Fällen dürften die Gleichheitsrelation und die ordnende Relation einen realen Sinn haben, obwohl dieser selten klar ausgedrückt wird. Die ordnende Relation, worauf die reihenartige Ordnung beruht, gründet sich oft direkt auf die Leistung in der Testsituation; z.B. kann diejenige Versuchsperson zu einer „höheren“ Stelle der Reihe gehören, die mehrere Testprobleme als eine andere gelöst hat oder die mehrere solche Testprobleme gelöst hat, die von weniger Versuchspersonen gelöst worden sind, oder die die Testsituation mit weniger oder kleineren Fehlern löst. Dann wird auch die Gleichheitsrelation einen empirischen Gehalt besitzen, und die Definitionen dieser Relationen sind meistens von solcher Form, dass die gehörigen logischen Forderungen erfüllt werden. In solchen Fällen steht eine quantitative Ordnung auf korrekter Grundlage.

Eine quantitative Ordnung ist aber für die Messbarkeit nicht hinreichend. Noch eine weitere Relation muss notwendig in der Erfahrung gefunden werden, eine Relation, die der Relation der Distanzengleichheit logisch gleichwertig ist. Bei keiner psychologischen Messung aber können wir eine solche Relation in der Erfahrung finden, nämlich eine Relation frei von Zufälligkeit und Künstlichkeit.

Eine Ausnahme bilden vielleicht die Messungen der sog. Sinnesqualitäten, u.a. die Messungen der erlebten Stärke des Tones und des Lichtes, denn die erforderlichen quantitativen Relationen können direkt der Erfahrung entnommen werden. Unter günstigen Umständen können wir nämlich erlebnismässig, unmittelbar, entscheiden, ob zwei Töne gleich stark sind bzw. welcher der stärkere ist. Weiter können wir alle Unterschiedsschwellen als gleich gross ansehen; eine andere Möglichkeit



wäre es, erlebnismässig zu entscheiden, wann die Stärke eines Tones genau in der Mitte zwischen den Stärken zweier anderer Töne liegt. Wenn empirisch festgestellt worden ist, dass diese Relationssysteme den gehörigen logischen Forderungen genügen, so kann die Messung der Intensitäten auf beiden Grundlagen durchgeführt werden <sup>2)</sup>. Wir können jedoch feststellen, dass der psychologische Wert solcher Messungen sehr gering ist; Bedeutung haben sie eigentlich nur für die Physiologie.

## VI

Unter den Psychologen scheint die Auffassung ziemlich allgemein zu sein, dass die Massskala eine richtige sei und man die richtigen Zahlenverhältnisse erhalten würde, wenn die Verteilung eine sog. normale Verteilung ist, die dem bekannten Gauss'schen Gesetz entspricht. Es lässt sich nicht leugnen, dass die normale Verteilung einen Spezialfall bildet, da sie gewissermassen die einfachste von allen möglichen „zufälligen“ Verteilungen ist. Wenn man auf die normale Verteilung einen so grossen Wert beilegt, so liegt dahinter der Gedanke, dass man auf diese Weise die einzige richtige Messmethode finden könnte und die Garantie dafür erhielte, dass das Mass an jeder Stelle der Skala „wirklich“ gleich gross sei. Wie wir schon festgestellt haben, ist dieser Gedanke unrichtig; es gibt kein Kriterium, wodurch man entscheiden könnte, welche Massskala die richtige sei.

Die normale Verteilung ist allerdings insofern bedeutungsvoll, dass sie die „Masszahlen“ eindeutig bestimmt. Wollte man durch die Zahlen nur die quantitative Ordnung der Elemente darstellen, so könnte das System der Masszahlen in sehr verschiedenen Weisen verändert werden; man könnte sozusagen die Skala beliebig „verdichten“ oder „sperrern“. Da man durch die Forderung der normalen Verteilung eindeutig determinierte Masszahlen erhalten wird, so ist es angemessen zu untersuchen, welch

<sup>2)</sup> Es sei bemerkt, dass diese zwei Möglichkeiten zwei verschiedene Quantitäten definieren. Die reihenartige Ordnung wird in beiden Fällen dieselbe sein, nicht aber die Distanzengleichheit (d.h. vielleicht empirisch dieselbe, aber nicht notwendig). Diesen zwei Möglichkeiten entsprechen die bekannten Gesetze von Weber und Fechner, die also voneinander logisch unabhängig sind, was auch gelegentlich behauptet worden ist.

einen empirischen Inhalt die so erhaltenen Masszahlen besitzen mögen.

Die normale Verteilung, die neben der Grössenordnung der Elemente ein wesentlicher Faktor ist, hängt von der Klasse der Elemente ab, von der Population, wie man in der Psychologie zu sagen pflegt. Wir stellen fest, dass wenn zu der Population ein Element hinzugefügt oder davon ausgeschlossen wird, so wird die Verteilung verändert; falls die Verteilung auch weiterhin als normal angesehen wird, so werden die Masszahlen sich ändern. Die Masszahlen sind also von der Population wesentlich abhängig. Da die Population immer zufällig ist, so sind auch die davon abhängenden Masszahlen zufällig (man vergleiche hier die Möglichkeit, die Temperaturskala auf die Forderung der normalen Verteilung aller beobachteten Temperaturen zu gründen!). Da die Masszahlen in der Psychologie von zufälligem Charakter sind, ist es kein Wunder dass die Psychologie kein einziges mathematisches Gesetz vorlegen kann.

## VII

Der Unterschied zwischen den exakten Naturwissenschaften einerseits und der Biologie und Psychologie andererseits besteht nicht darin, dass etwa die Eigenschaften der leblosen Natur messbar seien, die der belebten Natur dagegen unmessbar. Wir haben festgestellt, dass derartige Redeweisen inhaltslos sind und dass man eigentlich jede Klasse von Elementen messbar nennen kann. Der Unterschied besteht darin, dass die Psychologie keine solche quantitative Relationsstruktur entdeckt hat, die auf wichtigen und zentralen Bestandteilen der Wirklichkeit beruhte. Ob solche Strukturen in den Objekten der Psychologie überhaupt zu finden sind, ist eine empirische Frage, die sich durch die Einführung neuer Berechnungskünste nicht lösen lässt.

### *Summary*

As the concept of measurable quantity is analysed it turns out that it is a purely logical one. The only characteristic common to all different kinds of measurements is the same logical structure, a relational structure, as it may be called its essential features being the empirical relations which determine the system of numbers attached to the measured objects as their measures.

If e.g. an equality relation, an ordering relation giving the linear order among the objects, and a quantitative equality relation of the distances formed by the objects are found, a quantity is completely determined and it does not have any additional empirical meaning. Empirically different basic relations constitute empirically different quantities. As a certain quantity is given if and only if these basic relations are given, the question about the right measurement of a quantity is without meaning.

The value of a method of measurement depends entirely on the value and importance of its basic relations. In this respect the fundamental physical measurements are of a very different nature than those used in psychology (although the difference is only a matter of degree): the former are connected with the most important properties of the reality while the latter are derived from relatively insignificant empirical relations; this holds true also for the statistical methods based on the assumption of normal distribution. Whether quantitative structures originating in central and important relations can be discovered within the field of psychology is an empirical question which cannot be solved by new computational methods.

## CHILD PSYCHOLOGY AND "CONTROVERSY OF SCHOOLS"

BY

HAIM ORMIAN (Jerusalem)

„.... combien il peut y avoir de diverses opinions touchant une même matière, qui soient soutenues par des gens doctes sans qu'il y en puisse avoir jamais plus d'une seule vraie...."

René Descartes, Discours de la Méthode

### Part I—THE CONTROVERSY

#### 1. MULTIPLICITY OF SCHOOLS AND ITS RESULTS

In the report of the 10th International Psychological Congress in Copenhagen in 1932, Burkhardt said: "The manifoldness of questioning, the polystratification of problems and the diversity of methods make the impression that the unity of psychology is still unsolved, and the crisis not yet overcome. Not even the subject-matter, 'the psychic', has a single meaning fixed and accepted" (15, p. 38). Or to quote Klüver from 1944: "An inquiry into what psychologists were doing about a decade ago revealed such a heterogeneity of aims, methods, and principles that all efforts to arrive at a satisfactory delineation of the field seemed futile. There is no doubt that a similar inquiry today would lead to essentially the same result" (28, p. 387). Even sharper is the criticism by Politzer of contemporary psychology (in 1929): "Its history for fifty years seems essentially to be one of an excess of criticism" (50, p. 19), but "the criticism is barren" (p. 33).

And if such was the state of matters in the thirties and forties, it was still more so in the early twenties of the 20th century. Every now and then a new psychological idea emerged which called itself "new", rightly or wrongly, and it pronounced a certain truth as *the* truth in the field of psychology.

Such was the situation, when there appeared the psychology founded by Dilthey and Spranger, which called itself cultural science or understanding psychology, and which set its principles



out of a strongly negative attitude both theoretical and methodological towards a psychology which they called natural science or explanatory psychology. A similar situation occurred when Freud declared his theory based entirely on the unconscious as the only origin of human behavior, in contradiction to the psychology before him, which dealt with conscious experience only. That was the same situation, when functionalists asked "What the mind is for?" instead of "What the mind is?" in the manner of the existentialists (structuralists). The "objective" psychology denied the right of inner experience to be subject-matter for experimentation. The Gestalt psychology denied strongly "atomism" and accepted only the "wholeness" or "structure" as well in molecular phenomena (perceiving, learning) as in the molar behavior of the living organism. It would be possible to bring more antitheses as evidence for a dialectical development of psychology in the last 60 years.

It is easy to understand that the representatives of these theories, emerging out of a storm of opposition, and attacking each other aggressively, could not come to an understanding on any subject. The approaches turned quickly to be "schools", and the schools—various "psychologies".

Ragsdale speaks about six primarily noteworthy schools (5, p. 46 s), while Heidebreder analyses the approach of "Seven Psychologies", seeing first of all the American psychology, both "original" and influenced by Europeans (24). The number is sometimes reduced to five schools, as by Touroff (69) or Woodworth in the 1931 edition of the "Contemporary Schools..." (76—but ten in the 1948 edition). It happens that authors enumerate as many as 23 (Messer 37) or 16 trends (Müller-Freienfels 40). The beginning student may get entirely confused, when coming across such a variety of views, as is to be seen in Flügel's "A Hundred Years of Psychology" (20). We may easily understand Driesch's claim that some relevant psychological problems "have in our days reached a critical point", first of all "the fundamental materials and laws of normal psychology" (17, p. IX).

What is now the respect in which they differ? It is not only the heterogeneity of *interpretation of facts*, a thing well known in scientific research, but the relevance of the phenomena themselves. And that is the main point! Messer speaks about 14 trends (out of 23) which differ from each other in their atti-

tudes towards the *subject-matter* of psychology. There is not the question—how to understand the knee-reflex, the fear to remain alone in the room, or the child's obstinacy, and their connections with the organism as a whole or with the social life. The bone of contention, which precedes each interpretation, is: What does psychology deal with? Does it deal only with conscious processes, or solely with unconscious processes, with unconditioned and conditioned reflexes, only with the personal structure etc. Thus, the psychologists have not come to any common point or to any united idea even on the fundamental question, what is the subject-matter of psychology—not to speak about methods of research. "In fact, the heterogeneity of certain fields of inquiry in 'psychology' is so great that the unity often appears to be established by simply employing the *word* psychology.... We find treatises and text-books which seem to be as different as books on physical chemistry and pedagogy or books on astronomy and gastronomy", as Klüver said (28, p. 398, 391). No wonder that some psychologists (indeed, *only* German-writing) claimed that "the choice of a psychological method should.... depend on what kind of man one is" (70, p. 6). In any case, "Unity is certainly the most urgent need of psychology"—stressed Politzer (50, p. 15).

Moreover, the representatives of certain schools have revealed their theories with a great deal of zealotry and aggressiveness. "Gestalt psychology—asserts Keller—has been accused of denying its ancestors and ignoring its contemporaries" (27, p. 96), but, of course, each psychological "school" or "movement" has been denied by the followers of other "systems".<sup>1</sup> Each one would question the value of the scientific work of all its antagonists. There were schools more aggressive and intolerant, like the behaviorists and the psychanalysts, or moderate and tolerant, like the "dynamic" approach of Woodworth.

And the result—confusion! That is what Müller-Freienfels said in the early thirties: "Each system leaves important problems without solution.... but each system has its special advantages"; nevertheless *it is impossible* to unite them to-day

<sup>1</sup> David Hume: "Nothing is more usual and more natural for those, who pretend to discover any thing new to the world in philosophy and the sciences, than to insinuate the praises of their own systems, by decrying all those, which have been advanced before them" (quoted from 27, p. 79).

into one system. Luckily it's a wonder that "in spite of the impetuous dispute, psychology did not dissolve in an entire chaos". There is no possibility of compromise, at least not in the near future, as the differences between the schools are an unavoidable result of necessity. Finally, he demands from the psychologist one thing only—to see not only the weak points of his opponent, but also his advantages (40, pp. 136, 139).

The confusion among general readers, which arose out of such a situation, is well known. But much worse was the destructive impact on professional educators and on parents. They got intoxicated with concepts and subconcepts like complex, acquisition of habits, "männlicher Protest", personality, unconsciousness, level of aspiration, or adjustment, and they believed that they educated in accordance with those concepts.

The striking controversy between the schools was to be seen especially in two fields: (i) in the whole of the German psychology; (ii) in the U.S.A., where there was a tendency to exclusiveness, especially in the orthodox behavioristic psychology.<sup>2</sup> English, Swiss, French and Italian psychologists, also representatives of smaller countries, were much less engaged in the quarrels between schools.

We have to face an entirely different situation in the U.S.S.R. The history of the Soviet psychology is a history of the endeavour to see mind and psychology in the light of dialectical materialism. The position of pure and applied psychology, ontological, epistemological and methodological problems have been solved from the marxist standpoint, not without the influence of the party. Soviet psychologists call to cut off all the connections with the Western idealistic psychology. A strong opposition is felt against Gestalt psychology because of its subjectivism, against psychoanalysis because of its individualism and decadence, and against psychotechnics and testology because of their pseudo-scientific and fatalistic ways. Soviet psychologists stress the social aspect of behavior, the unity of social and personal, the idea of mental development, the unity of consciousness and action, the plasticity of the mind and the unity of theory and practice. Lenin's "theory of reflection" is the basis of all deliberations, also the endeavour to see marxist psychology as a part of a dialectical system of science, and beyond the Western controversies and schools. But all these efforts have not yet created a comprehensive system of dialectical psychology (cf. 68, 33, 42 b).

---

<sup>2</sup> That is what J. B. Watson said: "Behaviorism is new wine that cannot be poured into old bottles" (72, p. 41).

Also in West Europe there have been before the World War II and afterwards certain psychologists, who have analysed the situation of psychology and its crisis from the standpoint of dialectical materialism<sup>3</sup>. Some prove the justness of one of the analytic "schools" in the light of Marxism, e.g. as Rühle and other German socialists in favour of the Adlerian psychology, or as Fromm, Politzer, Bernfeld and Reich in favour of psychoanalysis. In opposition to them, some psychologists reject, as marxists, all the analytic trends, e.g. Schwarz (cf. 55).

In Germany there were schools which showed more or less the inclination to take into consideration achievements of their "opponents". The Würzburg School e.g. acquired the experiment from natural psychology and the description of experience from cultural psychology. They, too, were inclined to apply problems and accomplishments of Gestalt psychology and of the personalistic approach.

But by no means was this the case with the heads of humanistic psychology, of psychoanalysis or individual psychology, and partly of the adherents of the Gestalt psychology. That is what Spranger says: "What is the advantage of knowing that in puberty changes occur in the body of the adolescent?"; "We cannot understand the soul and its functions by means of knowledge of the body". And something struggles in us against statistics (61, pp. 22, 45 etc.). So also Bernfeld in his "Die heutige Psychologie der Pubertät" (7), which totally condemns five German "Psychologies of Adolescence" because of the negative or indifferent relation to the psychoanalysis of each of the authors. That is the *only* scale for Bernfeld to estimate the scientific value of a psychological book.

## 2. WAY OUT OF THE SITUATION

Little by little the absurdity of the situation made itself felt. Things had gone so far that one psychologist simply did not understand the language of his colleague, i.e. he did not understand the terms, nor did he try to do it. But on the other hand the situation led some circles to endeavour to find a connecting bridge.

---

<sup>3</sup> The opposition that is at the bottom of the 'crisis' of psychology lies in the opposition between the *idealistic* and the *materialistic* psychology" (50, p. 129).



In the course of time two methods were seen trying to find a way out of the confused situation. The first way was rather that of German psychologists, whereas the second was followed by psychologists in English speaking countries; but it was the second way which was successful.

The ideas of Karl Bühler, as summarised in his profound "Die Krise der Psychologie", may serve as an instructive trial to overcome the confusion of psychology which finally produces psychologies. Bühler put forward the main question in a form derived from Kant: "How is psychology possible?" He discusses and compares the main principles of several schools—the introspective psychology, the behavioristic, humanistic and analytic, and comes to the final statement: There is *one* psychic life, but it may be considered from three aspects—the experience, the meaningful behavior, and their relation to mental formations. "I state the thesis that each one of these three aspects is possible, and no one is dispensable in the one science of psychology" (14, p. 29).

This is a typical example of overcoming definitions and statements by means of new definitions and statements. That was what many of German psychologists did as they tried to find a common language with their opponents by clarifying basic concepts. Their number was not great, but their scientific weight was heavy enough, e.g. the "Theoretische Psychologie" by Lindworsky (32, Einleitung, § 2).

At the same time there appeared in Germany, but first of all in the U.S.A., other trials to overcome the quarrel of schools. Some did it by extreme empiric ways, saying: Let us not pay attention to the fight of principles; let us deal with facts. Some tried another way: Let us see the facts first, and this will lead us to the forming of principles.

The excellent textbook edited by Boring *et al.* deals with facts and questions which had been shared by various schools—sensations, responses, motivation, consciousness, individual differences, and the ego. The authors describe phenomena, seek their roots, find out the connections between them and make use largely of the ideas proclaimed by representatives of different principal views. Only in the 1948 edition of this book are schools mentioned (21½ pages), but primarily to state at last: "What was

good in all the schools is now simply part of psychology" (11, p. 11).<sup>4</sup>

In the most popular textbook, the "Psychology" of Woodworth (now of Woodworth & Marquis), a caution is felt to touch the differences of principles. In the first chapter, however, the authors state when experiment, introspection or other methods should be used. There are data we can observe objectively, and there are subjective data "furnished by the subject from self-observation" (77, p. 16). The reader will easily understand that this follows consequently from the nature of the investigated data. The authors use as well achievements of introspective psychologists (e.g. of the Würzburg School), as of the psychoanalysts (e.g. the concept of identification) and the Gestaltists. A single time only are the "Schools" mentioned (not in the last edition!), when the author states: "There is no reason why *the outsider* (italics mine, H. O.) should not find value for himself in the work of all the schools" (l.c., edit. 1945, p. 583). Woodworth's middle way is evident even in his "Dynamic Psychology" devoted first of all to *his own* approach to psychological problems. The outlook on the modern psychological movement and the description of problems and methods within the contemporary trends show Woodworth's willingness to stress the common traits and not the opposing ones (75, ch. II—III).

As another example of such a moderate way may serve the modern textbook by Munn. The author does not deal with schools; he states that psychology is "the science of experience and behavior" (41, p. 16) and connects by this simple statement two kinds of phenomena, the relation of which created in modern psychology a complicated and seemingly unsolvable situation.

Most writers endeavour to explain their methods when writing monographs, not textbooks. And here are some examples in chronological order:

In Boring's paper "Psychology for Eclectics" is felt a strong direct and indirect opposition against the onesidedness and esotericism of "scholastics". His method is based on *facts*, and

---

<sup>4</sup> How this author is far from the school controversy, we can learn from the last sentence in his review of the 1948 edition of Woodworth's "Contemporary Schools...": "I venture to predict that in his third edition in 1965 he will go so far as to leave the word 'Schools' out of the title" (J. Abn. Soc. Ps. 1949, 44, p. 281).

he is not inclined to be enthusiastic about the truth of "schools". That is why "The eclectic of 1930 will accept both behavior and phenomena as the data of his psychology".<sup>5</sup> He "will certainly wish to take the view of Gestalt psychology.... choose modern structuralism.... choose formally to include the functional interest in his psychology", but first of all "he hurries back to his laboratory to start new research" (42 a, pp. 119—127).

Another example may be "A Briefer General Psychology" by Murphy. Here are some quotations from ch. XXIV devoted to the Psychological "Schools" (inverted commas in original!): "One is often bewildered by popular books and articles presenting some theory or 'school' of psychology. In these popular accounts one notes a tendency to oversimplify the subject matter of psychology. The unwary reader is left with the conviction that there is just *one* coherent, straightforward system of theories to which *all* psychological facts must conform" (43, p. 517). And after he discussed briefly some foundations of behaviorism, existential and Gestalt psychology, he comes to the following conclusions *based on facts described in his book*: "Each of these schools has something important to contribute, but none is the last word". What have the "schools" in common? The use of experimental methods and of quantitative devices (l.c., p. 528 s). *Method* may be the favourable factor to draw the schools near to each other. This American optimistic approach becomes strikingly clearer in comparison with the aforementioned wavering German ideas of Müller-Freienfels. And both the books were published in the thirties!

Woodworth too, hardly mentions "schools" in the numerous editions of his textbooks, whereas in his "Contemporary Schools of Psychology" he argues from a typically empirical point of view: "If the existentialist presents a good analysis of heat sensations, or of colour experiences, we accept it with thanks. If the behaviourist shows by experiments on little children how conditioned fears may arise, we are free to use that finding in our own psychology. If the Gestalt psychologist should show that all learning depended on some degree of insight, we should

---

<sup>5</sup> Similarly in 1948: "Psychology deals with both the *behavior* of man as it appears in his responses and with *consciousness* as he finds it in his immediate experience"; this statement involves also the unconsciousness (11, p. 4 s).



revise our conceptions of learning accordingly. If the purposivist convinces us that the individual is never passive when a stimulus reaches him, that is another important point to be dealt with. If the psychoanalyst opens our eyes to the importance of sex motivation, we thank him for that." A psychologist without prejudices should have such a relation to each actually positive achievement wherever it comes from. Which school is it that will win finally? "All and none" (76, edit. 1931, p. 227, 231). The truth is in "The Middle of the Road". "Every school is good, though no one is good enough. No one of them has the full vision of the psychology of the future" (edit. 1948, p. 255).

In a similar way speaks Flügel: "There can be little doubt that each school has developed certain important aspects of psychological truth; but, if this is so, it follows that no one school possesses the sole key to truth. . . . In this way there may come into being one 'psychology' with many methods, in place of the several 'psychologies' that exist to-day" (20, p. 359 s).

Also Touroff (an American psychologist writing Hebrew) warns against the danger of one-sidedness in searching for mental facts. He argues thus: "Our acknowledgment of the great importance of libido function in man's life. . . . does not imply the belief in the *sole* rule of the libido". Such is the case when we accept "the existence of physiological responses to artificial stimuli. . . ., of strong natural tendencies called 'instincts'. . . ., of the right to solve psychological problems by paying attention to the personality as wholeness". We are not bound by these "acknowledgements" to accept the rule of any principle in reference to our *whole* behavior (69, p. 239 s).

This is the way of psychology in English speaking countries—in consequence of the empirical approach of philosophy and science in these centres. Sometimes such a way is to be found also in German speaking countries, even already in the early twenties of our century. Here are three examples: Bernfeld the orthodox-minded psychoanalyst, Koffka one of the founders of the Gestalt psychology, and W. Stern the founder of the personalistic trend.

Bernfeld never went a step beyond the psychoanalytical doctrine. Nevertheless he is ready to admit that there are some advantages in methods unused by analysts. It is clearly stated in his "Psychologie des Säuglings" (6), and also in "Die Psycho-

analyse in der Jugendforschung" (5), which is a methodological introduction to the book devoted to social psychology of the youth. Bernfeld was the only German psychoanalyst who tried to find common points with other psychologies; he tried to find a way to the Gestaltists (8), or e.g. general, pre-analytic sources for Freud's concepts (9).

Experiment and measurement, external and inner observation, investigation of children's creations (poems, stories, diaries, letters)—all these seem to him acceptable techniques and materials *behind* the psychoanalytic interpretation of children's and adolescents' behavior, "although it is clear that the use of psychoanalysis enables a decisive progress in the understanding of any [psychological] problems.... It is very difficult—he asserts—to have the impression that the analytic publications are based on empirics, on strict observation and on cautious formulating of relations, because all these means of guarding against autistic thinking, called scientific method, which are added to the description of results of investigation, as a rule don't appear in the analytic investigations" (5, p. 8 ss). Applying "exact methods", used in psychology, analysts will put an end to these methodological principal objections used by their opponents. Bernfeld himself does it in his: "Vom Gemeinschaftsleben der Jugend" and in other papers, which contain many exact description of methods and of material—a thing rather rare in psychoanalytical researches.

Also Kurt Lewin was of the opinion that "The attempt of the psychoanalysts to base general laws entirely on case studies and therapeutical work seemed methodologically unsound to most scientists" (31, p. 3). This methodological disadvantage encourages the opponents of psychoanalysis to claim that its principles are not sufficiently based on exact methodology. Flügel, too, lays stress on this disadvantage in his historical book (10, e.g. pp. 287, 293), and still more in his systematic paper which appeared in Murchison's "Psychologies of 1940": The psychoanalytic "methods are, however, still highly cumbersome and inconvenient; it has, in fact, not yet reached the experimental stage" (42 c, p. 393). Also Murphy stated in 1929: "The impossibility of experimental or statistical control of complicated factors unearthed by the intricate and arduous process of psychoanalysis" is one of the factors of the critical relation to this view (44, p. 334). Heidebreder: "The technique of psychoanalysis is dismissed by many as unsound for scientific investigation" (24, p. 422). And, last not least, Sears in 1942 (according 3, p. 35) and in 1944 (56, p. 306): "The subjective character of the system has proved a difficulty, .... since it does not permit the conventional scientific tests of proof and disproof." So also Ellis in 1949 (18, 124) and 1950 (19, p. 189 ss). And, of course, there have

been made many serious attempts to adjust experimental methods and techniques to psychoanalysis (cf. Sears 56; Wolff 74).

This sharp external criticism of psychoanalytic *methods* began in the later twenties and in the thirties (Murphy—edit. 1929; Heidebreder; Lewin). The internal criticism began in the thirties (especially Flügel and Horney), and it was to be felt more clearly in the forties (e.g. Sears; Alexander-French; Wolff; Ellis). But some change, even in the fundamental *approach*, *although indirectly*, was to be seen with Bernfeld still in the twenties. E.g.: "There is a task before psychoanalysis, to assert itself as psychology—even if not as something complete and final (7, p. 4). All other psychological methods and psychoanalysis "don't by any means exclude each other, but they complete each other, since Freud's psychology studied intensively just the psychology of drives, which was too much neglected by other schools" (6, p. V). The one requires the research and the achievements of the second. It is not by chance that Bernfeld does not deal in this book with perception, language, or intelligence, but he leaves them out, being aware of this fact. And we should read with interest those chapters written by a representative of this school, who never tried to search subject matters, as perception, learning, thinking or development of language.

Koffka's "Grundlagen der psychischen Entwicklung" (29) may serve as a second proof of the tendency in Germany to find a common language. The biological outlook, structural standpoint, and behavioristic methodology were mingled in Koffka's theory, and they are helpful to each other in the investigation of mental growth in the first years of human growth. Parenthetically: During the 12th Congress of the German Psychological Society (spring 1931) Koffka said in his address on "Die Bedeutung des Behaviorismus für die vergleichende Psychologie" that he appreciated behaviorism as long as it is a method of research and as a working hypothesis; but he firmly denied its claim to serve as a Weltanschauung. Furthermore, behaviorism cannot be accepted as the one and unique explanation of man's behavior (cf. 45, p. 139).

This claim of behaviorists that they, only they, are in possession of the key to the understanding of mental behavior, aroused the firm opposition of



MacDougall, who called himself the "Arch-Behaviorist" rather than Watson himself. It was he, McDougall, who fought against the theory that psychology means knowledge of conscious states only, and demanded the objective investigation of behavior "two classes of data both useful and indispensable for the *one science* (italics mine H.O.) of human nature properly called 'psychology'" (2, p. 53, 58).

W. Stern in his "Allgemeine Psychologie" laid stress on connection and contact between psychologies—as his point of view for a long time. "The personalistic hypothesis does not exclude other theories and points of view (except the purely mechanistic), but bears a constructive relationship to them; and although it is homogenous, onesidedness is avoided" (63, p. VIII).

Briefly: The opposition against onesidedness began to appear also in German psychology. It might have shown a rather moderate attitude towards psychological schools (at a slower pace than in the U.S.A.), were it not for political factors which put an end to German psychology in 1933. The "race theory" started its terroristic rule also in the field of psychology, any freedom of ideas was abolished, and every scientist who was not a sworn Nazi was sacked immediately. In consequence of such a state, a dead silence was felt also in the field of psychological school controversy. It does not mean that the controversy was settled, but the "schools" were destroyed, and their adherents were silenced. The only papers allowed to be published were such as Jaensch's pamphlet about the importance of psychology in the "German movement" (26), Peterman's "The Problem of the Race-soul" (49), "papers" on "Hitler as political psychologist", and more futile "scientific papers" written in the spirit of Gleichschaltung. German psychologists denied psychoanalysis or individual psychology because non-Germans were founders of these "sickly" theories, whereas the German wholesome Volk would not "accept" them by any means—according to Kroh in his "Die Aufgabe der pädagogischen Psychologie und ihre Stellung in der Gegenwart" (30, p. 321 s).<sup>6</sup>

<sup>6</sup> Fr. Baumgarten-Tramer made an important contribution to the analysis of the German psychology and psychologists since World War I — attitudes towards war, Volk, Hitler as leader and "political psychologist", and their influence on the German Geist before and during the World War II (4). Cf. also the comprehensive study by Weinreich of "Hitler's Professors" (73).

### 3. IN ENGLISH SPEAKING COUNTRIES

In the U.S.A. we may follow, even before the thirties, a development of methodology and principles based primarily on facts, and only in the second line on principles. Murphy pointed out already in 1929 that in the British and American psychiatry "much eclecticism. . . prevails", and general and developmental psychology likewise approached certain psychoanalytical findings and views (44, p. 334).

After 1933, the activity of American psychology was influenced, no doubt, also by the immigration of a number of distinguished German and Australian psychologists, who contributed, consciously and unconsciously, to a compromise. First of all, the adherents of Gestalt psychology gained authority, since their theory had been known even before, owing to translations. Also the Freudians and Adlerians got adherents, and quite unexpectedly the share of introspection rose. This situation, however, did not strengthen the extremists, and the "immigrating" ideas melted away with the previous American ones. To Boring, "During the 1930's the isms pretty well dropped out of psychology" (11, p. 11), until Sargent could say in 1944: "At present no active school exist" (54, p. 4). This formulation is, no doubt, exaggerated. There did not appear, indeed, *new* schools, and the tendency of the majority of psychologists to co-operation has grown, instead of the trend to be under care of any separatistic system. "Academic psychologists"—as Murphy asserts in 1949—"might at times protest against eclecticism. . . . But [the contrary is true of] those engaged in diagnosis and in therapy" (44, p. 419).

Fractures appeared even among psychanalysts. As a characteristic proof may serve the careful attitude of Susan Isaacs in her well known books on child psychology. The author, who had been dealing for years with psychological and pedagogical problems in the famous Malting House, discussed some psychological problems avoided or even rejected (as problems!) by psychanalysts; e.g. mental tests, norms and standards, motor ability and reasoning. And perhaps it is not by chance that she used to speak about "psychologist", and not "psychanalyst".

Not only representatives of psychanalysis tried to draw nearer to psychology, but the same way can be observed in the ironically so called "academic psychology", which started going



towards psychoanalysis. Signs of this approach may be found in Bonaventura's "La Psicoanalisi", first edition of which appeared in 1938. The author dealt with psychoanalysis as a psychologist (10, p. 29) and he tried to bring out from their esoteric loneliness terms and concepts which had been the sole "possession" of psychoanalysis. He succeeds in finding common points in researches, which psychoanalysis claimed as "its own", e.g. unconsciousness, sexuality, dream or neurosis (l.c., ch. II, V-VII). It turned out that there are common problems, that methods may be instructive to each other, and that even the results do not differ as much as is taken for granted by extreme "ideologists" (espec. ch. X).

That is what brought McDougall nearer to psycho-analysis. "My criticism [against psychoanalysis] is ruthless, it is nevertheless entirely friendly; and it aspires to be constructive" (36, p. VI). He wanted to reach a "fusion or synthesis" of his own instinct theory and that of Freud, as both of them refer, after all, to the same subject. In his opinion, Flügel may serve as the connecting bridge, because he appears "to be, in fact, a Freudian, possibly the only Freudian, capable of entering into fruitful controversy with psychologists" (l.c., p. 24).

This remark is not exact enough, as is made clear by what was formerly said about Bernfeld and Isaacs. But it is true, that the author of "A Hundred Years of Psychology" knows to evaluate in the right way each psychological trend. Moreover, several years before the publication of his historical treatise, Flügel claims in his paper on "Psychoanalysis" "that the barrier between psycho-analysis and other psychological systems is being slowly broken down" (42 c, p. 375). In the "Psychologies of 1925" no psycho-analytic paper was published; unlike in "Psychologies of 1930". Flügel clears up what are the difficulties and the points requiring correction to establish a "closer relation between the psychoanalyst and the 'academic' psychologist. In this matter, questions of method are of supreme importance" (l.c., p. 393).

Let us return to McDougall. He endeavours to find parallel fundamental principles in his hormic psychology and in psycho-analysis, e.g. their firm opposition to behaviorism. He explains, too, what are the essential differences, and in which points there are merely differences of terms. The "ego" of Freud is called by him "character", whereas the "superego" corresponds to his "moral character" etc. (36, p. 104).<sup>7</sup> The author

<sup>7</sup> As to the psychological terms: The confusion and the differences of

discusses broadly (ch. III) various changes in psychoanalytic terms and concepts, which have arisen for the last years, and he realizes that there is a possibility for psychoanalysis and psychology being drawn nearer to each other (i.e., espec. ch. II).

We see that serious and fruitful attempts were made by psychologists (not psychoanalytic) and by psychoanalysts to learn from each other.<sup>8</sup> These attempts may prove useful; they may destroy the unnecessary Chinese Wall separating the two groups of psychologists. This partition was perhaps understandable (not acceptable!) in the early days of psycho-analysis, but is by no means so nowadays. The two sides may take advantage of each other, provided that they cease to be "sides" in the former sense of the word.

#### 4. THE SITUATION TO-DAY AND ITS FACTORS

After World War II, there is to be found not only a factual approach of the schools towards each other, but also a steadily increasing recognition of the necessity of such a fact. Two examples: Thorpe admits that "The point of view of 'Child Psychology and Development' is one of broad organismic development which accepts the contributions of more than one school of psychology. This position might well be called *patterned eclecticism*" (66, p. VII). Much more instructive may be the Harvard University Report about "The Place of Psychology in an Ideal University". Some statements in the Report are a proof of fact that this idea is becoming wide-spread and almost an "official" attitude. In the chapter on "Observation and Interpretation" it is stated that a syllabus of psychology should necessarily contain

---

their meanings have been a real obstacle to the development of psychology. This is the case when two schools give two different names to the same phenomenon (e.g. learning — conditioning), or when they call different things by the same name, e.g. psychology (!), observation, drive, motive, instinct. As Tumarkin said: "If two styles of thinking are so far apart, then the identity of expression is rather confusing than clarifying" (70, p. 2 s). Only so is clear Watson's famous claim, from his first "behavioristic" paper (Ps. Rev. 1913), against the terms used in psychology and his intention to cancel them throughout, or at least to adjust them to the behavioristic concepts (71, e.g. p. 24, 276, 297 ss — German edit.). If he succeeded, it would still strengthen the split between behaviorists and other psychologists.

<sup>8</sup> Cf. the summary of Snyder about the psychotherapeutic counseling (59, espec. pp. 322—333).

various points of views (= schools), and this "will stimulate investigations to test the more significant of its formulations" (23, p. 9).

In short—we are now further from the *onesided extreme approach* and rather near the *middle-way*. It will not be an exaggeration to see in such a book as "Current Trends in Psychology" (16) a clear and convincing sign that the old style of the controversy has died away. The volume deals with psychology as profession, child psychology, personnel psychology, human engineering, but it does not mention "schools" and their controversies. We find, instead, investigation of human life, subjects arising from the needs of every day. This fact is significant for the practical attitude of American psychology almost from its beginning. Titchener's statement that "The attitude of science . . . is before all things a *disinterested* attitude" (67, p. 42), i.e. that pure psychology is not attached to the needs of real life, is rather an exception than a rule in that country.

And now we have to ask the fundamental question: Why has the storm calmed down—at any rate in certain centres?

The answer must include several subjects.

In the U.S.A. psychology has been marked by the preference for empirical investigation and learning facts to the speculative treatment of principles since its beginning (cf. 2) up to our day (cf. 47). The reference is to practical psychologists only, when Klüver exclaims "that most psychologists in this country seem to be too much occupied with their experimental problems to worry about a definition of psychology or to find time and interest to reach an agreement as to what aims a science of psychology should pursue" (28, p. 393). The just mentioned "learning of facts" was not lacking in "points of view" or summarizing laws. But the elasticity of these "points of view" was large, and, as we see, they knew, in most cases, to fit themselves to the facts.

In the U.S.A. even psychanalysts show to-day signs of being less dogmatic (or "orthodox"), and inclined to adopt methods used in general psychology. Besides the proofs mentioned above (Wolf, Sears), let us give two other representative examples: (a) The Chicago Institute for Psychoanalysis, which has searched for shorter and more efficient means of psychotherapeutic treatment, states that their "main contribution to the growth of psychotherapy is a return from stereotyped thinking to experi-



mentalism" (1, p. 24). (b) Ellis analysed all the papers which appeared in the first half of 1947 in four leading American psychoanalytic journals, and did not find an "undue concern.... with some of the simplest and most accepted requisites of scientific research". But on the other hand he found a serious improvement of psychoanalytic research by approach, raising of exploratory hypotheses, avoidance of over-generalization, cautious interpretation and forthright self-criticism (18, p. 124, 138 ss).

German psychologists, however, even experimentally-minded, had laid, from Wundt on, much more stress on the philosophical basis. This tendency in certain papers appeared in the form of proper speculative reasoning, as in the school of Dilthey-Spranger, and to a certain degree in the papers of such psychologists as Brentano, Külpe, Stern, Bühler or Lindworsky. Even to-day we find this German tendency to discuss the situation of psychology and its problems among Swiss German-writing psychologists, e.g. with Suter (64, espec. Einleitung and ch. 7). This does not imply, by any means, that German psychologists did not deal with empirical facts. But they stressed much more, or at least—more, the through principles rather than through approach particular investigations<sup>9</sup>.

Two outstanding examples of the value of abstract principles even in the German applied psychology: (a) Hylla tried to explain in his "Testprüfungen der Intelligenz", why American testology was spreading quickly, while the German progressed by slow steps. The German psychologists—he asserted—hesitated to begin with an extensive use of intelligence tests as long as they were not clear enough about their theoretical fundamentals. The Americans, however, "with their strong sense of reality and opposition to theoretical subtlety", paid attention to the real needs and they developed intelligence testing as a method of examination, even without a "theory". And the result—the development of American mental tests and testing (25, p. 100). (b) Moede, too, almost apologized that he and his colleagues began in 1917 testing gifted children in Berlin, although they

<sup>9</sup> W. Stern was one of the few German psychologists who firmly required the investigation of "totality and multiplicity of traits", or investigation "von oben unten" and vice versa, according to his personalistic view (63, p. 30 and ch. I/III).



hardly had a theory concerning intelligence and its testing (38, p. 146 s).

Moreover, "scholastics", tending mainly to principles, subjugated the mere *choosing* of psychic phenomena for their research work to the need and possibility of proving *their* principal approach. Thus, psychanalysts turned, first of all, to drives, natural science psychologists to sensations, cultural science psychologists to the "meaningful objective formations", and the Würzburg school to the processes of thinking. Hence the collision between the principles and their far-reaching results, principles which were not elastic enough, because psychologists chose a *certain* group of mental facts for research (reasoning, personality, libido etc.) and handled them successfully, each in his own way and according to his principal Einstellung, but the validity of such principles could not be more general, although "scholastics" pretended to give laws and principles of *general* validity.

Finally, the needs of real life became the deciding factor—education, personnel work, industry, army, or therapeutic counseling. The hasty prediction of Titchener proved unreal, that it would be impossible, for perhaps a hundred of years, to apply psychological findings. There were very real needs that showed decidedly the right way and they undermined the basis of the old style school controversy that lately had lacked stability from within.

It is worth mentioning that in Europe after World War II the recognition has increased that psychology is a practical, applicable science, e.g. Révész in "Die Bedeutung der Psychologie für die Wissenschaft, für die Praxis und die akademische Ausbildung der Psychologen". He was dealing with this problem not only because psychology is the fundamental science of the cultural sciences (52, p. 17) and of the natural sciences (p. 58 ss), but mainly because of "the rising interest in psychology and the widespread enlargement of its fields of application" (p. 5).

## Part II—THE CONTRIBUTION OF CHILD PSYCHOLOGY

### 5. EMPIRIC AND GENETIC APPROACH OF CHILD PSYCHOLOGY

Observation of facts and not speculation about principles on the one side, and fruitful stimulation of the applying aspect on the other side—these two tendencies are more outstanding when

we deal with mental behavior of the growing organism. *The developmental psychology may show the way, how to strengthen the foundations of psychological investigation*, while overcoming the difficulties of both true and imaginary controversies and contradictions between its principles.

The developmental psychology may be the most constructive factor in the effort to relieve psychology of the barren discussion about abstract principles. It will show how to tread on the right way to investigation, without neglecting principles. It achieves these principles, corrects them, extends or limits their range within the process of investigation itself.

This branch of psychology deserves to play this part owing to both its subject-matter and its methods.

First of all: Numerous are the experiences of this psychological branch, but few are the investigators who dared to search this field armed with *general principles only* without the burden of "little" systematic and detailed investigations of *facts*. Spranger's "Psychologie des Jugendalters", born in the circle of the German psychology in the early twenties of our century, is one of the rare examples within the developmental psychology of a rather deductive way (61, espec. ch. I).

It does not require special proof that the lack of exact empirical methods must lead to a mainly speculative manner and to rather literary expressions. Experience has shown that descriptions of such kinds are lacking in general validity, as they do not reflect the true mental reality, only giving statements derived from other statements—even in case there are *also* certain proper observations. Besides, evaluations necessarily caused by such a manner of treatment of psychological data dim the picture. And convincing is the example of Spranger's above-mentioned book.

Only this empirical caution makes us explain certain sentences in J. B. Morgan's Introduction to the 3rd edition of his "Child Psychology": "When the first edition of this book was published in 1931, it was difficult to find enough experimental material for an integrated text.... Where gaps in the experimental material were found, it was necessary to fill these with hypotheses and speculative material". A different situation was seen in the 2nd edition of 1934, as many new studies were being published in the meantime, and still more in the 1947 edition (39, p. VII). Indeed, even in 1931 there were empirical investiga-

tions in certain languages carried out with methodological carefulness. Anyhow, Morgan's words express a strong spirit of a developmental psychologist's tending towards a true empiricism.

There is no contrast between Morgan's statement referring to the situation in the thirties and those of Lewin published in the same period. It might seem that the latter says something contrary when he states: "A number of branches of psychology have reached a stage which makes their unification increasingly urgent. Child psychology, for instance, has collected a great number of facts about speech, play and other forms of behavior at different age levels". Hence the need of a comprehensive theory (31, p. 4). But there is no contradiction between the two authors, if I do understand their statements given directly and indirectly. There is a difference only in emphasis. Both of them oppose theories based on speculative reasoning, both demand a theory based on facts and enabling the investigator to understand the majority of the investigated phenomena.

To say to-day "There is no psychology but the empirical" is, indeed, nothing new, and a proof from psychological *writings* is not convincing enough. We have to learn still more by observation of the *object of our investigation*, i.e. the developing child.

The investigation into the child's actual behavior in his first years of life shows that the old-fashioned controversies between exclusive hereditarianism and exclusive environmentalism may find their way into an archive of antiques. The modern version of the controversy "environment-heredity" or "nature-nurture" is based on the concrete investigation of the growing child and the conditions of his growing.<sup>10</sup> As to important contributions it is sufficient to mention here the stories of children who grew among animals, the famous story of Gue and Donald, and, above all, investigations of identical twins and foster children.

Such investigations enable us to see in a proper light not only the general psychological question of the mutual relation between "nature" and "nurture". They contribute to a great extent to the understanding of mental development. All differences of opinion referring to the *particulars* of this subject

---

<sup>10</sup> An interesting proof of investigation of the relationship between the outlook of scientists on "The Nature-Nurture Controversy" and their attitude toward social, political and economic questions was given recently by N. Pastore (see 48).



(e.g. the famous Iowa controversies) do not change the basic fact that development is a process of *interaction* between what is called "nature" and what is called "nurture", i.e. all psychic changes occurring in the growing human organism are results of the mutual action and constant "struggle" between the organism and his environment.

From another standpoint: Methods and treatment of biology (investigating the "hereditary equipment of man") and of sociology (investigating the "environmental conditions of growing") are both called to co-operate in the investigation of the growing organism because of the nature of the mental development. There is a clear line of development in the process of this interaction which may contribute to the understanding of our principal problem dealt with. Changes in behavior, beginning with the human fetus up to the last stages of the social development in adolescence, are a history of transition from the objective preponderance of the "nature" to the objective preponderance of the "society", i.e. also from the biological view and method to the sociological one. Parenthetically: Contemporary psychology inclines to stress more and more the "nurture" component in the process of interaction of nature and nurture. Pastore's abovementioned book demonstrates it doubtlessly, in particular the comparison of Terman's attitude in 1932 and in 1948, as given in an instructive passage of his letter written in 1948 to Pastore: "I still strongly suspect the existence of race differences, but I am now inclined to think that they may be less than I formerly believed them to be" (48, p. 88).

This state of transition in childhood from "nature" to "society" enables us to overcome the historical duality (or opposition) between the natural science psychology and the cultural science or social science one. Child psychology is neither biology nor sociology, although it is most closely connected with both of them, following the methods of general psychology. Although tied to biology and sociology with respect to the general outlook and ways of research, and learning from them—it does not cease to be psychological science with its own problem and methods.

## 6. OUTLOOKS AND METHODS

The actual quality of the connection between psychology and biology on the one side, and between psychology and sociology



on the other is made clear step by step as we constantly observe the process of the child's development.

We find in the "controversy of schools" echoes of tendencies of philosophers in favour of the autonomy of mental life, or of the mind as a function of the body or a certain psychophysical theory. The empiric-minded psychologist, however, will push aside the controversy on this field, and he will translate the abstract problem of "mind and body" into the language of concrete phenomena seen both by psychologist and biologist—each of them by means of his methods. The connection between the two aspects of the behavior of the organism may be illustrated best by actual facts, showing that there is here no abstract relation between "body" and "mind", but real connections between the various patterns of the behavior of the organism.

And here are several examples: The correspondence between the development of the sense organs and the sensations, between stages of the growth of height and weight and general mental development, between the activity of the thymus, thyroid and sexual glands during later childhood and early puberty on the one side and such phenomena as the need of movement, sexual dreams, development of libido or changes in the relations between girls and boys in those stages of development on the other side—all these are actual "psychophysical" problems well known in child psychology.

Investigation of such phenomena, rising before us when dealing with development of mental behavior, may serve as the best way for finding a strong basis for discussing the relation between psychology and biology. A similar clearing up requires the relation between psychology and sociology, as I tried to demonstrate by investigating an actual problem of the social psychology of adolescence in "Adolescents' autobiographies as a psychological source" (46, p. 68).

Furthermore, by observing mental development, we may find a way to the solution of another question always connected with the fundamental approach to mental phenomena—methods of investigation. Child psychology has influenced to a great extent the principal view and has enlarged the *methodological* horizon of psychology in general. We have learned that there exist no fruitful and efficient studies of mental behavior, but those using the development approach. Since mental development

is both a biological and a sociological phenomenon, the methods used will be necessarily biological *and* sociological. This is not a logical consequence from the above assumptions, but an actual state of affairs. As a proof may serve the comprehensive paper of Anderson about "Methods of Child Psychology" in Carmichael's Manual (3).

What are biological methods? Various forms of experiment, continuous observation, time sampling observation for finding an "inventory" of behavior, observation of groups of children (with a control group or without it) in order to study e.g. their collective games.

Child psychology as well as the animal one taught psychologists to stress objective behavior, and to use appropriate methods. But the use of behavioristic, objective *methods* does not, by any means, compel us to interpret each mode of child's behavior by the mechanistic formula: S—R, since the observation of children's behavior shows that there are spontaneous, purposeful movements, also conscious processes, which are accessible to us. Thus, in the first month of children's life we will use behavioristic, "objective" methods only; also during the whole of childhood we will not divert our attention from that kind of method. But in the course of children's development we come across such phenomena as imagination, language or reasoning where pure "objective" biological methods must fail. So the biological methods must be enlarged, when such mental experiences come into consideration, and some new method must be added to the objective ones, used throughout the whole of childhood.

The environment as a factor in the child's development caused the need of using methods and techniques to investigate not the "pure" maturation (if there exists at all "immanent" laws of development), but the contribution of the environment to the process of mental growth. For this purpose a type of tests was constructed measuring the individual's general development (cf. 22; 13) by a "developmental quotient", and taking into account also the contents derived by the child from his environment.

Moreover, in later childhood children are able to express something of their conscious experiences. They may report more or less, according to the age and to the investigated problem, but they are able to carry out partial introspection. Thus, a child

psychologist will use this method which brings him nearer to the child's conscious processes, such as thinking, perceiving or imagining.

An investigator of attitudes or interests in vocation, or in school subjects, in national, political or social problems, interest in fiction, poetry or moving pictures etc., will not find any more efficient and more comprehensive method than a questionnaire taken in psychology from the social sciences. It is a simple, systematic method, and repeatable in the same conditions. Questionnaires were administered to thousands of children in various centres, and the method proves useful, whatever the difficulties might be. Nowadays, questionnaires are no longer enthusiastically and uncritically administered, as was the case in the time of St. Hall and his followers in the U.S.A., and in the first quarter of our century in almost all cultural countries, when the study of "ideals" was so much "en vogue", following E. M. Darrah-Dyke and the "Child Study Movement" in the U.S.A. Nevertheless, after the decline of enthusiasm for the questionnaires, we have learned to use them methodically and in an efficient and more technical way (e.g. 34, ch. VII), and this technique became a useful tool regarding the study of later childhood and adolescence.

Here is another problem demonstrating that observation of the child may add another contribution to the solution of the discussed principal question: A child felt aches in his hand which had been cured some month before. An X-ray examination did not show any organic basis. Thus, there is a psychic factor of the renewal of these aches. We cannot but turn to a way which may detect unconscious impulses caused by something which happened in the child's environment. The investigation of the unconscious contents will necessarily lead us to an method of investigation the nucleus of which will be psychanalytic. "Psychanalytic" does not mean exclusively the "orthodox" way, requiring a rather intuitive relation to the patient, but may also include experimental methods. It is certainly not a matter of chance that the first attempt on the part of psychoanalysts to apply experiments, investigating such behavior situations as fixation, regression, frustration or aggression, were carried out by means of experiments on children or animals (cf. 56, espec. p. 307 ss). Similarly Wolff asserts that we may improve our



understanding of "the depth of the child's personality" by means of the experimental method (74, p. 296). He also realized his methodological demands, as described in his "Personality of the Pre-school Child".

And a similar methodological problem: The widespread and well known impudence of boys 11—14 years of age, their desire for independence, their common claims towards adults and their oversensitiveness—don't all these phenomena teach us that certain unconscious impulses bring forward this opposite rebellious conduct?! Their conduct is not a sign of self-confidence and strength. It proves just the opposite—influence of weakness, rising out of a misty feeling: "Here I am, grown up, and they treat me as if I were an infant. Let me show them what I am able to do." To understand such phenomena shouldn't we turn to certain notions and explanations which have grown up in A. Adler's climate?!

The analysis of the situation from the standpoint of the relations between pedagogics and psychology will throw additional light on the problem. In spite of the controversy of schools, many of them may be helpful in education. Let us take several examples: When a teacher is in need of a psychological basis, when coming across questions of learning at school and at home, he certainly will turn first of all to the laws of learning as adapted by representatives of Gestalt psychology; he who deals with problems of sexual education may make use also of psychoanalytic knowledge of sexual development in childhood; modern educational psychology is based on the American functionalistic approach; the emphasizing of early childhood both by the behaviorists and psychoanalysts led to the emphasis on need of organized education in the pre-school age, etc.

That is not the end of patterns of behavior, most frequent in children's life, each of which requires an approach and a method corresponding with its nature. But on the other hand: "All these various approaches—as Wolff rightly stressed—must be checked and balanced against each other, if an objective and at the same time full understanding of the child is to be attained", and "great misconceptions" in child psychology are to be avoided. Indeed, the history of modern child psychology shows that there is a danger of one-sidedness in stressing one psychic factor and negating other, while using one approach alone and denying the other (74, p. 293 s).



In conclusion: Child psychology shows once more that *factual contrasts between "schools" are much less felt than the abstract controversies.*

Furthermore, child psychology is not only a "pure" science. The immense development of this branch in the U.S.A. has shown its practical value—in spite of Titchener's famous opinion that it will be impossible to apply psychological findings perhaps in the next hundred years. Child psychology remains an applied branch of psychology too, however emphatic we may be in conceiving the attitude that child psychology arose as an applied science and that it exists for the sake of its applicability (16, p. 50). Whatever the history of child study may be, it is clear nowadays that the practical aspect is of great importance.

In other words: Child psychology is that branch of psychology which is most closely connected with the interaction of the so-called "pure" and so-called "applied" psychologies. It is doubtless an exaggeration when J. B. Morgan states "The most important problem is not to study the abstract nature of his [the child's] mind, . . . but to teach him the significance of life and some habits and attitudes" (39, p. 5), or according to the first sentence in the "Child Development" by Breckenridge and Vincent: "The main reason for studying child development is to improve the lives of children" (12, p. III). We can, by no means, comply with the theory that child psychology is nothing but "*ancilla paedagogiae*". Nevertheless, the manysided applicability of child psychology in pedagogics is a very important phenomenon.

Psychologists analysing educational activity from a psychological standpoint, or (in other terms) applying psychological achievements within educational theory and practice, show the right way to the psychologists, though they are yet "on war-terms". From this standpoint we will review two textbooks of educational psychology, of the most widespread, and we will see what the authors state.

Skinner, editor of "Educational Psychology" and author of its first chapter ("The Nature and Scope of Educational Psychology"), deals merely with "methods used in psychology" (58, p. 13); "schools" don't appear, not even as problem to solve.

To quote the second author, Gates, who mentions explicitly the matter of psychological "systems": His book and his colleagues' "is non-sectarian" and it "does not ardently espouse any one

system of beliefs"; it "includes facts, principles, and applications in a form acceptable to persons with preferences for any one of the major systems". The extremists only, as Gates stresses, had exaggerated the differences among the psychological systems. But *real research work must* bring psychologists nearer to the real truth, that in proportion depends very little on the principles of *one* system (21, p. 9 s).

Professional and non-professional educators come across children's conduct which presents educational problems, sometimes difficult to solve. Hence they are bound to learn the ways of adjustment and maladjustment of the individual and their factors. Pedagogical counsellors and guides are supposed to give advice regarding the choice of a course of study in keeping with the abilities, aptitudes and interests of the child. In some cases they have to advise the proper educational way desired from the point of view of the individual and society, to avoid some difficulties of adjustment deviations. In such situations they have to know the whole life of the child in his environment—to be able to analyse his situation and to suggest a reliable advice to his educators.

Consequently we need a method which will enable us to comprise these factors, and we find it by paying attention to the real character of the situation, i.e. the many-sidedness of the phenomena. *One* method that arose on the ground of *one* psychological doctrine will be, by no means, sufficient. In fact it will be a *group* of methods, satisfying the demands of a biologist, a physician, a sociologist and a psychologist at the same time. Investigation of environment, psychometric and projective tests, personality inventories, free planned observation, general medical and neurological examinations—all these together form the famous clinical method which arose out of child psychology on the one side and of educational and adjustment needs on the other.

From the standpoint of the general problem of this paper it is important to stress that the clinical methods are in no way connected with *one* single psychological doctrine. They are connected with *various* views, and consequently compounded of various methods. It corresponds to what Snyder stresses in his critical summary, especially when treating counselling dealing with children: "No single method of therapy is applicable to all problems. . . . There has arisen recently a rather strong develop-

ment towards eclecticism" (59, p. 363, 367; cf. also 44, edit. 1949, p. 420).

We find this eclectic attitude also in other branches of applied child psychology—in vocational guidance and in psychological service. In this fields the pure and the applied view go hand in hand.

We cannot imagine an efficient psychological or vocational guidance, except one based on pure research of child's interests, abilities and aptitudes and their distribution among children of various ages, of both sexes, and of various environmental conditions. The chief aim of testing child's intelligence, musical aptitude, or vocational interest is to enable us to find his position among children of his age, sex, educational level and socio-cultural position. This is possible, when the investigating of *many* children precedes the testing of the *individual*. The investigation gives us norms to appraise the ability of each examined individual, and this enables us to fix his position within the group he belongs to. And this is the required knowledge for any guidance in any field. It is just the common development of applied child psychology and the present boom in clinical psychology which requires more seriously a "pure" basis for the practical activity <sup>11</sup>.

On the other hand, the need of work in psychotechnical laboratories and in pedagogical and psychological clinics has encouraged "pure" research in various centres. One fact from the history of child psychology may serve as an example of research resulting from the applied approach to child psychology. I am referring to the famous scale of Binet. It was practical needs that led him and Simon to search for "questions" (= tests) to examine the level of intelligence of children, supposed to be mentally retarded. These practical examinations, however, caused an immense movement in child psychology, and as result the test

<sup>11</sup> How convincing is in light of these facts Rosenzweig's claim, expressed in the Symposium devoted to "Clinical Practice and Personality Theory", when he agrees with Fernberger's statement: "Pure science must come before applications". But, unfortunately, "the theories of the academic psychologists and the practices of the clinical psychologists have for the most part developed independently of one other" (65, p. 4). I.e. that laying stress on applied psychology (applied child psychology included) does by no means imply the abolishment of pure (child) psychology; just the opposite.



became also a "pure" instrument of research, as may be seen e.g. in the studies of genius by Terman and his collaborators. The main subject and aim of those researches is a "pure" problem—development of gifted men from childhood to adulthood measured also by means of mental tests.

As we see, child psychology shows that the discrimination between pure and applied psychology is not in keeping with theoretical principles, nor does it suit the needs of life. In this respect, American psychologists went so far that they used the term "psychologist" in reference to a person dealing with practical psychology, such as a personnel worker, clinician or vocational counselor (cf. 16).

Case study method and mental test on one side, experiment and questionnaire on the other, bring forward once more the question which has been since Wundt one of the main bones of contention in psychology. The question is: What does psychology tend to, what does it aim at—to discover laws referring to the psychic behavior of all, or to give descriptions of individual cases? This point was at the bottom the nucleus of one of the controversies between the natural science or explanatory psychology and the cultural science or understanding one. This controversy is only an echo of the systematic and methodological controversy between *Naturwissenschaften* and *Geisteswissenschaften* in Germany, or to be put into the epistemological language of those days (Rickert, *Windelband*) a difference, leading to opposition, between nomothetic and ideographic sciences.

The nature of developing man and the quality of the science dealing with this nature show us beyond any doubt that we need in child psychology nomothetic methods as well as the ideographic ones. The actual choice in each case depends mainly on our scientific purpose. Sometimes we are interested in investigating many children to find psychological rules. Each method chosen in this situation in accordance with the investigated mental function (introspection, objective observation, experiment, questionnaire) must be of a nomothetic character. Our purpose is to find out general or differential *laws*. But often we shall examine an individual case, mostly for a practical purpose<sup>12</sup>. Each method

<sup>12</sup> The terms "to research" and "to examine" have been used here following W. Stern's explained discrimination between "Forschungsexperiment" and "Prüfungsexperiment" (62, ch. V—VI).



chosen in such a situation in accordance with the respective practical aim (vocational or pedagogical guidance, psychological service, personnel work) will be based on the principles of the ideograph method, our purpose being to examine a concrete individual. It refers to each kind of test, of personality inventory, of projective tests, life-histories or of the wholeness of these techniques, namely the case study method.

This is one more contribution of the developmental psychology to the question discussed now: There is no principal or methodological opposition between the use of "nomothetic" ways in investigating many persons with the view to discover psychological laws, and the use of "ideographic" ways in examining individuals in order to find a manysided picture of them. But in psychology, investigation of rules or laws must precede examination of individuals.

Let us now turn to an additional general psychological problem, the solution of which may be derived from the situation of the whole child psychology. The problem is: What is the subject matter of psychological research—the personality or its "parts", the mental wholeness or its "single" ways of behavior and features? As a matter of fact, the question is nowadays by far not so acute as it used to be. Strong attack against "atomism" has no real basis in contemporary psychology. Even if nowadays a psychologist deals with a partial way of behavior, he takes into consideration only "parts" (in inverted commas!) wishing to lay stress upon his opposition to the opinion that there might exist independent features of ways of behavior. Nevertheless, the question of actual *preference* may still be asked: Should the investigation of the behavior of the whole organism be preferred or that of its "parts"? On this point, child psychology may bring an additional considerable contribution to the settling of unnecessary and harmful controversy between psychological "schools".

In what circumstances does this question arise within child psychology?

If we face educational difficulties, it is clear that difficulty lies in the molar person, and not in one of his "molecular" features. Careful examination of an individual may reveal feeble-mindedness, deviations in attention, sexual disturbances, overdeveloped self-reliance, bed-wetting or any other "partial" factor. But the

aim of such a diagnosis is not to find the factor of the deficiency in a *single field of behavior*, but the basis of maladjustment of the *whole individual*. The diagnosis only fixes the focus of the individual's maladjustment. In consequence of the intra-individual relations between all kinds of features and ways of behavior, the *whole personality of the child is affected*, and not only a "single field". The extent of this affectation depends on the "centricity" or "periphericity" of the deviating feature.

Consequently, there is here no contrast either between personality and single features or ways of behavior. After all, it is the personality which is acting, not its dependent "parts". Sometimes we have in view rather the wholeness, the acting person, without distinguishing its "parts", sometimes rather some special way of behavior draws our attention. When is the second way allowed? Only when we bear in mind that the "particular" is but a part of the wholeness, i.e. that it reveals the whole acting organism.

This is an additional contribution of child psychology to the general controversy, both with regard to the principal approach (the organism as acting molar pattern) and to the methodological one (research comes before examination).

## 7. SUMMARY AND CONCLUSIONS

It is unnecessary to continue and to bring additional facts. I hope it has been made clear enough that the development of the child, the phenomena which are being revealed in it step by step, as well as factors of these phenomena, show us that *both looking at the child from one point of view and using one method only, may perhaps reveal one part of the truth, but the whole truth will remain unclear*.

The above mentioned examples, taken from different fields of child psychology, not only require the use of various approaches and methods within child psychology itself, but still more they pave the way for the whole of contemporary psychology.

Here are the summarized conclusions based on the above discussion:

(a) Our principal views on mental behavior as well as on aims and methods of psychology, as science describing this behavior and finding and systematizing its laws, depend also on our social and economic needs.

(b) Certain problems which once filled psychologists with enthusiasm and served as the basis of main theories, are no longer of any use. Who will nowadays devote his studies to psychophysical questions in the manner of Fechner and Weber?! How much less attention is now paid to the research of sensations, even in reference to the infant?!

(c) Child psychology (but not only it) shows how to overcome seeming contrasts and controversies, which have been growing up in modern psychology. First of all, we have learned to overcome the historical ambiguity of the biological vs. sociological view on man's behavior. We have learned also to overcome what seemed to be a contrast between seeing personality and its "parts". Nevertheless, it is clear nowadays that the importance and the efficiency of a sociological-minded attitude has been growing in psychology, and that the personality comes before its "parts". Investigation of man in social situations has become one of the fundamental problems in modern psychology.

(d) As to the controversy between the static and dynamic view, there are the facts that decide. The research of developmental age shows that the dynamic view is the right one, as it is to be learned e.g. by observation of the impact of impulses upon man's behavior or of the influence of the determining tendency upon thinking processes. Nevertheless, there are cases when the static approach is advisable, e.g. when we want to compare groups of children of various mental levels or to fit the position of an individual within his group (cross sectional method).

(e) Child psychology abolishes also the seeming contrast between a pure and applied approach to man's behavior. Accordingly, the corresponding seeming contrast between the discovery of psychological laws and the examination of individuals as aim of psychologist's activity are diminished.

(f) Furthermore, objective examination of facts leads to the relevant conclusion that there is no contrast between "subjective" experience and "objective" behavior.

(g) The research of children shows us, much more than investigation of adults, the position of the unconscious in an individual's life. On the other hand, however, we are neither minded to neglect the investigation of conscious processes nor to see them as second-rate elements in mental life and, consequently, in psychological research.



(h) "Methods and techniques do not fall into mid-Victorian classifications of right and wrong, of all and nothing" (3, p. 38). The mental behavior is recognizable only by means of *various* approaches or views and methods. The choice of a fitting suitable method, however, and the interpretation of the findings do not depend on chance or on arbitrary decision. It is the quality of the mental facts that shows every time the methodological way. Investigation of each developmental stage requires fitting of methods and interpretation, as does each phenomenon—provided that we learn to fit in our view and method to the actual material, and not let it be the consequence of our general outlook on the "nature" of mental behavior.

(i) The psychologist is not satisfied with the question "What is going on", i.e. with recording and describing. As a scientist he wants much more. Description goes hand in hand with interpretation—"why" (cause), "what for" (purpose), and "in connection with" (correlation). Clearly, psychologists are not of one mind as to the interpretation itself. It should be a naïve opinion to say that disagreements and controversies within child psychology (and within each other branch of psychology) will or can disappear. But the main thing is that controversies of opinions should be based on a respectable relation to facts, without neglecting principles.

(j) That is true—psychology is still a science in the making. Nevertheless, we are able to observe recently a growing tendency towards harmony, stability, and unification, instead of a multiplicity of isolated "one and unique" systems, and of the past spirit of zealous rivalry or at least emotional discussion among them. So the modern psychologist has overcome to a certain extent inner differences of opinion, true as well as imaginary, and he adapts to himself *various* balanced views, methods and techniques. But, at last, a simplified eclectic compromise can not agree. We have to endeavour a more fundamental solution.

For orthodox-minded "scholastics" this does not present great difficulties. In their opinion the problem has not existed. It is easy to state (less easy to demonstrate!) that there exists *one* centre, *one* phenomenon or *one* feature which explains the *whole* behavior of man, may be—reflex, libido, personality, Gestalt, psychic formations, endeavour to power, instinct, adjustment or learning (in the history of psychology also—association). Indeed,



it is true that "Psychology needs—following Lewin—concepts which can be applied not merely to the facts of a single field like: child psychology, animal psychology, or psychopathology, but which are equally applicable to all of them", yet not one single concept such as reflex or totality (31, p. 6 s). James complained that psychology has "not a simple law in the sense in which physics shows us laws", and it is consequently "only the hope of a science" (quoted from 60, p. 3).

Nevertheless it is very doubtful, whether there exists any *central single phenomenon or a way of behavior* which can be seen as a focus of the whole behavior of man. Hence, it is doubtful whether there may be found one sole principle, or even a group of principles (similar to the fundamental laws of classic physics) which may serve as a unique explanatory "key" of man's entire conduct.

The concept of the Middle ages, stating that man is a "unitas multiplex", seems nowadays rather insoluble in its first part—"unitas", whereas modern psychology is more able to overcome the difficulties of "multiplex". Anyhow, known phenomena and methods don't yet allow us to become acquainted with the "unity" in a systematic and empirical way. Only he who is satisfied with the rationalistic approach only or with empiricism comprising a certain limited area of research, may find a satisfactory answer in to-day's state of psychological searching. Psychology is still waiting for a system of concepts and principles which will enable us to explain man's mental life from a broader aspect comprising all ways of psychic behavior in all branches of psychology.

Some psychologists have tried to find an irreducible minimum of necessary rules. Spearman suggested the laws of association, retentivity, control, constant output, fatigue etc. (60, p. 20). Suter suggests a few principles—of reality, subjectivity, relative awareness, structure and dynamism (64, p. 160—170). Keller hopes that the gestaltistic, structuralistic, functionalistic, and behavioristic approaches will become a point of view that "will be known, not as a *system*, but as *psychology*" (27, p. 98). Ruckmick endeavours to find, by help of other psychologists, "the basic principles of psychology as a whole". He suggests, for his part, principal postulates which are common, in his opinion, to all psychologists, i.e. province of experience as the subject matter of psychology,

possibility of analysis, and borrowing of "related concepts and facts from other affiliated sciences" (53, p. 325 ss).

The number of suggestions is not few, and not all important ones have just been mentioned. In this situation we need research, planned research (cf. 35; 57) of facts (without the ironical adjective—"so called" facts) which will show us, whether we are able to reach in this generation a satisfactory system of principles that comprehend the whole range of psychic behavior. To quote Klüver: "We venture the prediction that a Copernican revolution initiating an era of rapid progress in psychology, if it ever occurs, will be brought about by the monographic treatment of specific problems and the detailed analysis of particular phenomena" (28, p. 405 s).

For my part, I hope that child psychology is the right branch of psychological science to be able to contribute its important share to finding a solution that will not be a philosophic, pedagogic, or sociological, but a psychological one—in spite of the aid we need on the part of these sciences.

#### BIBLIOGRAPHY

1. Alexander, F. & French, T. M., *Psychoanalytic Therapy* 1946.
2. Allport, G. W., *Die theoretischen Hauptströmungen der amerikanischen Psychologie der Gegenwart*, Z. Pd. Ps. 1924, 25.
3. Anderson, J. E., *Methods of Child Psychology*, in Carmichael L. (ed.), *Manual of Child Psychology* 1946, 1—42.
4. Baumgarten-Tramer F., *German psychologists and recent events*, J. Abn. Soc. Ps. 1948, 43, 452—465 (German 1949).
5. Bernfeld, S., *Die Psychoanalyse in der Jugendforschung*, in *Vom Gemeinschaftsleben der Jugend* 1922, 1—11.
6. ———, *Die Psychologie des Säuglings* 1925 (Engl. transl. 1929).
7. ———, *Die heutige Psychologie der Pubertät. Zur Kritik ihrer Wissenschaftlichkeit*, *Imago* 1927, 13, 1—56.
8. ———, *Die Gestalttheorie*, *Imago* 1934, 20, 32—78.
9. ———, *Freud's earliest theories and the school of Helmholtz*, *Psychoanal. Quart* 1944, 13, 341—362.
10. Bonaventura, E., *La psicoanalisi* 1950.
11. Boring, E. G. et al. (ed.), *Foundations of Psychology* 1948.
12. Breckenridge, M. E. & Vincent E. L., *Child Development* 1950.
13. Bühler, Ch. & Hetzer, H., *Testing Children's Development from Birth to School Age* 1935 (German-Kleinkindertest).
14. Bühler, K., *Die Krise der Psychologie* 1927.
15. Burkhardt, H., *Bericht über den X. Internationalen Kongress für Psychologie*, Z. Pd. Ps. 1933, 37—42.

16. Dennis, W. (ed), *Current Trends in Psychology* 1947.
17. Driesch, H., *The Crisis in Psychology* 1925 (German 1926).
18. Ellis, A., Towards the improvement of psychoanalytic research, *Psychoanal. Rev.* 1949, 36, 123—143.
19. ———, An introduction to the principles of scientific psychoanalysis, *Genet. Ps. Monog.* 1950, 41, 147—212.
20. Flügel, J. C., *A Hundred Years of Psychology* 1935.
21. Gates, A. I. et al., *Educational Psychology* 1949.
22. Gesell, A. et al., *The First Five Years of Life* 1940.
23. Gregg et al., *The Place of Psychology in an Ideal University* 1947.
24. Heidebreder, E., *Seven Psychologies* 1933.
25. Hylla, E., *Testprüfungen der Intelligenz* 1927.
26. Jaensch, E., *Die Lage und die Aufgaben der Psychologie. Ihre Sendung in der deutschen Bewegung und an der Kulturwende* 1933.
27. Keller, P. S., *The Definition of Psychology* 1937.
28. Klüver, H., *Psychology at the beginning of World War II*, *J. Ps.* 1949, 28, 383—410.
29. Koffka, K., *Die Grundlagen der psychischen Entwicklung* 1925 (Engl. transl. 1928).
30. Kroh, O., *Die Aufgabe der pädagogischen Psychologie und ihre Stellung in der Gegenwart*, *Z. Pd. Ps.* 1933, 34, 305—327.
31. Lewin, K., *Principles of Topological Psychology* 1936.
32. Lindworsky, J., *Theoretische Psychologie im Umriss* 1932.
33. London, I. D., *A historical survey of psychology in the Soviet Union*, *Ps. Bu.* 1949, 46, 241—277.
34. Lundberg, G. A., *Social Research* 1942.
35. Marquis, D. G., *Research planning at the frontiers of science*, *Amer. Ps.-ist* 1948, 3, 430—438.
36. McDougall, W., *Psychoanalysis and Social Psychology* 1936 (Germ. transl. 1947).
37. Messer, A., *Einführung in die Psychologie und die psychologischen Richtungen der Gegenwart* 1927.
38. Moede, W., *Berliner Begabenschulen* 1919.
39. Morgan, J. B., *Child Psychology* 1947.
40. Müller-Freienfels, R., *Die Hauptrichtungen der gegenwärtigen Psychologie* 1931.
41. Munn, N. L., *Psychology. The Fundamental of Human Adjustment* 1947.
42. Murchison, C. (ed.), *Psychologies of 1930*, 1930.
  - a. ch. 6: Boring E. G., *Psychology for eclectics*, 115—127;
  - b. part VI: *Russian psychologies*, 207—278;
  - c. ch. 20: Flügel, J. C., *Psychoanalysis*, 374—394.
43. Murphy, G., *A Briefer General Psychology* 1935.
44. ———, *An Historical Introduction to Modern Psychology* 1949.
45. Ormian, H., [12th Congress of the German Psychological Association] *Polsk Arch. Ps.* 1931, 4, 138—142 (Polish).
46. ———, [Adolescents' autobiographies], *Hahinukh* 1946/47, 20, 59—70 (Hebrew).

47. Pastore, N., Social influences upon psychological trends, *J. Gen. Ps.* 1948, 38, 15—29.
48. ———, *The Nature-Nurture Controversy* 1949.
49. Peterman, B., *Das Problem der Rassenseele* 1935.
50. Politzer, G., *La Crise de la Psychologie Contemporaine* 1947.
51. Ragsdale, C. D., *Modern Psychologies and Education* 1932.
52. Révész, G., *Die Bedeutung der Psychologie für die Wissenschaft, für die Praxis und die akademische Ausbildung der Psychologen* 1947.
53. Ruckmick, C. A., The principal postulates of psychology, *J. Gen. Ps.* 1938, 19, 325—331.
54. Sargent, S. S., *The Basic Teachings of the Great Psychologists* 1944.
55. Schwarz, Th., *Zur Kritik der Psychoanalyse* 1947.
56. Sears, R. R., Experimental analysis of psychoanalytical phenomena, in — Hunt J. McV. (ed.), *Personality and the Behavior Disorders* 1944, vol. I, 306—332.
57. Seidenfeld, M. A., Professional considerations in organized research, *Amer. Ps.-ist* 1949, 4, 410—414.
58. Skinner, Ch. E. (ed.), *Educational Psychology* 1947.
59. Snyder, W. U., The present status of psychotherapeutic counseling, *Ps. Bull.* 1947, 44, 297—386 (German transl. in *Psyche* 1949, 3, 401—459).
60. Spearman, C., *Psychology down the Ages*, vol. II, 1937.
61. Spranger, E., *Psychologie des Jugendalters* 1949.
62. Stern, W., *Differentielle Psychologie* 1921.
63. ———, *General Psychology from the Personalistic Standpoint* 1938 (German 1936).
64. Suter, J., *Psychologie. Grundlagen und Aufbau* 1942.
65. Symposium: Clinical Practice and Personality Theory, *J. Abn. Soc. Ps.* 1949, 44, 3—49.
66. Thorpe, L. P., *Child Psychology and Development* 1947.
67. Titchener, E. B., *Psychology: Science or technology*, *Pop. Sci. Month* 1914, 84.
68. Tomaszewski, T., [Principles of Psychology in the U.S.S.R.] 1949 (Polish).
69. Touroff, N., [The Psychology of To-day], vol. I, 1939 (Hebrew).
70. Tumarkin, A., *Die Methoden der psychologischen Forschung* 1929.
71. Watson, J. B., *Behaviorism* 1925 (German transl. 1930).
72. ——— & McDougall, W., *The Battle of Behaviorism* 1928.
73. Weinreich, M., *Hitler's Professors. The Part of Scholarship in Germany's Crimes against the Jewish People* 1946 (also Yiddish).
74. Wolff, W., *The Personality of the Pre-school Child* 1947.
75. Woodworth, R. S., *Dynamic Psychology* 1925.
76. ———, *Contemporary Schools of Psychology* 1948.
77. ——— & Marquis, D. G., *Psychology* 1947.



## THE STABILITY OF INTELLIGENCE TEST SCORES

*A report on some Swedish re-test investigations*

BY

TORSTEN HUSÉN, Stockholm

### THE CONCEPT OF STABILITY

In appraising the usefulness of a test, and in the majority of cases this implies the usefulness of an intelligence test, we have to face three basic problems. (1) To what extent does the instrument measure the same qualities as the criterion that one wishes to predict? This relationship will determine the predictive value of the instrument for behaviour which is independent of the test, i.e. its *validity* with regard to this behaviour. (2) With what degree of "accuracy" or "precision" does the instrument measure the behaviour which is relevant to the test situation. This precision is defined as the *coefficient of reliability*: which, determined by means of the split-half technique, indicates the influence on the test results, of random factors occurring in the test situation or in the test correction. (3) What is the degree of stability of the test scores obtained in a given test situation and with a given group of subjects? <sup>1)</sup> Stability can either be defined by means of the retest coefficient, obtained after a certain interval, or by the correlation between two tests designed to measure the same behaviour modality (intelligence tests, interest tests etc.) applied at certain given intervals.

The first two problems have received considerable attention in the methodological discussion of psychometric problems; while the third problem, the question of the stability of the test values, has not aroused as much interest, and the investigations dealing with it have been fewer in number. But in a large number of cases, where the long-range predictive value of a test is to be determined, the problem of stability cannot be ignored. This can be illustrated by the following example. One desires to predict, by means of intelligence tests, the ability of children

---

<sup>1)</sup> Cronbach (1949), p. 59 ff. distinguishes between a coefficient of equivalence (split-half-coefficient) and a coefficient of stability (retest-coefficient).

of 5—6 years of age to profit from the tuition given in the first class, or at a later stage, in the primary school. A test instrument with relatively high split-half reliability has been procured, let us say at  $+ .92$ <sup>2)</sup>. The correlation between the test and the teachers' ratings of the general ability of pupils in the fourth degree at 10—11 years of age is  $+ .45$ . This coefficient of validity is considerably lower than that usually obtained for older children with the criterion which is used here. One may then ask whether the low coefficient of validity is due to the fact that the test affords an "incomplete" measure of intelligence, or if the low validity is due to other circumstances. This question can only be answered by determining the stability of the test, i.e., how high the correlations are on retesting with identical or similar tests after 5—6 years. If it is then found, as in the present case, that the retest coefficient is  $+ .55$  after an interval of 5 years, this explains the low validity of the test. Actually, this cannot be higher than  $\sqrt{0.55}$ , i.e., 0.74.

Thus, it often occurs that we have a test with a high split-half reliability, but which, either in general, or in the case of subjects of a certain kind, age etc., measures a somewhat unstable behaviour. As a rule, when testing children for predictive purposes, it is of basic importance to ascertain whether the test instrument employed will, within a given temporal interval, put the persons measured in about the same rank order. If stability is low, the efforts expended to increase the validity of the instrument, for example, by including several more "relevant" items, may prove to be in vain. If, to apply this to the example given above, a test score obtained at 5—6 years of age does not prove sufficiently stable to have any predictive value for the test values obtained several years later, it will be necessary to investigate at what age a test should be administered in order to be adequately stable, and consequently also adequately predictive.

#### EARLIER INVESTIGATIONS

Investigations on the correlation between the two rank orders obtained by giving the same intelligence test at two test sessions, with different intervals between the sessions, form the

---

<sup>2)</sup> When referring to correlation coefficients, the naught in front of the decimal point has been omitted throughout.

empirical basis for any discussion of the "constancy of the I.Q." Surveys of the re-test coefficients obtained have been published, *inter alia*, by Hildreth (1926), R. L. Thorndike (1933, 1940), L. D. Anderson (1939), J. E. Andersson (1940), Bayley (1940), and Husén (1951) <sup>3)</sup>.

In view of the empirical material available, the following conclusions can be drawn:

1. The size of the re-test coefficient depends on the *length of the interval*. The longer the interval, the lower the coefficient.
2. The size of the re-test coefficient depends on the *age of the subject at the time of the first test administration*. Under conditions that are otherwise similar, *inter alia*, with the same interval, the coefficient is higher than that obtained in re-tests of older subjects. This principle holds good at least up to the age of 20.

Besides these two conditions, the re-test coefficient is affected by:

3. The reliability of the test, and
4. The systematic or random changes in the subjects during the interval between the two test administrations <sup>4)</sup>.

#### THE LENGTH OF THE INTERVAL

R. L. Thorndike (1940) has reviewed the re-test coefficients obtained, and has found that the relation between the interval and the re-test coefficient may be expressed in the form of an equation, where the re-test coefficient is the closest in linear relation to the interval. If a coefficient of  $+.90$  is obtained on immediate retesting, then after 10 months the coefficient will most probably be  $+.87$ ,  $+.81$  after 30 months, and  $+.70$  after 60 months. The relation established by Thorndike is valid for children under 10—12 years of age, after which time the re-test coefficients tend to decrease considerably as the interval is increased <sup>5)</sup>. Bayley (1940) has shown that a test with a relia-

<sup>3)</sup> In the Review of Educational Research, February 1941, February 1944 and February 1947; further material has been collected.

<sup>4)</sup> Cf. Husén (1950), p. 94 ff.

<sup>5)</sup> Freeman (1950), p. 503 gives  $+.90 - +.95$  for immediate retesting;  $+.85$  after 1 year,  $+.80$  after  $2\frac{1}{2}$  years,  $+.75 - +.80$  after 5 years, and  $+.78$  after 9 years. The decrease of re-test coefficients after 10—12 years will take place, when the second testing takes place after that age.

bility of  $+ .94$  for babies under 12 months, correlates  $+ .81$  with the scores obtained after an interval of 3 months, and practically 0 with the scores obtained after 6 years. Anastasi (1949) states that Gesell's schedule at 6 months, correlated  $+ .37$  with Merrill-Palmer scores of 2-year olds. Honzik (1938) found that the test result for children at 2 years of age, correlates about  $+ .30$  with re-tests obtained at 5—6 years of age. Bayley and Jones (1946) compared test scores at 10 years of age, on the one hand, and yearly results starting at one year of age, on the other. The series of coefficients obtained rises rapidly up to the age of five or six, when it reaches the level  $+ .80 - + .90$ , subsequently continuing to rise, but with rather low increments up to 10 years of age<sup>6</sup>).

The low correlations given here were at first considered to be partly due to the fact that "intelligence" during these early years was more dependent on environment than it is at a later stage, and partly to the fact that tests for successive ages cover rather different kinds of abilities. Both explanations are probably applicable. (see below.)

#### THE AGE AT THE FIRST TEST

In a paper published 1940, J. E. Anderson showed that *a priori*, correlations between tests, intervals remaining constant, may be expected to be the higher, the nearer the subjects approach adult age. It is simply a question of "overlapping" with regard to the number of common elements. The contents of a test for children of 9 has more in common with the contents of a test for children of 8, than with the contents of a test for children of 7. If a number of children of 9 years of age are tested, for example, with Terman-Merrill, then the majority of these will solve all the items up to and including those for children of 7. If these children are retested at 11 years of age, the majority will then solve all the items up to and including those for children of 9. If the scores for these children were previously obtained at 7 years of age, it will be found, that as a rule, they solved all items up to and including those for children of 5. Since the number of common elements must be greater when comparing the results

---

<sup>6</sup>) Jones (1946). p. 589.



of children of 9 and 11 years of age, the correlations must be higher here, than when comparing children of 7 and 9, where the number of common elements is less. By using, *inter alia*, random numbers, J. E. Anderson showed, that the correlation will increase when the number of common elements is greater, and furthermore, that the increase that was determined statistically coincided with the (empirically observed) coefficients. — As its age increases, the child becomes more and more “the father of the man”, since its actual capacity, as it grows older, embraces to an ever greater extent adult capacity.

The answer to the question, what the age of the child should be, for the test scores to be sufficiently stable to enable us to make predictions over a period of several years, depends on the degree of accuracy required by these predictions. R. L. Thorndike (1947) and Embree (1948) have shown that, when testing children of 9—11 years of age, nearly as good a forecast of school achievements at 18—20 years of age is obtained, as when testing them at 14—15. The curve for “forecasting efficiency” obtained by Jones (1946), shows that the tests carried out at the beginning of school attendance already display quite a high degree of stability. The practical consequences of these results from a pedagogical point of view are naturally of very considerable importance, as intelligence tests are continually being more widely used for purposes of differentiation.

In a follow-up investigation of 613 children, undertaken by the writer (Husén, 1951), where the children were tested at 10 years of age and then at 20, and in both cases with a group-test different in content, the correlation between the two series of scores was  $+ .72$ . On the first occasion the teachers rated the subjects' ability. The correlation between the rating and second test was  $+ .61$  (Husén, 1950). Although in the second case there was an interval of 10 years between the rating and the test, no significant decrease was noticeable in the correlation. Apart from the presumably higher reliability of the second group test, the estimation of ability, by means of tests or ratings, possesses quite a high predictive value, although these were carried out already at the age of 9—10, i.e., before the pupils were differentiated by some of them leaving to receive higher education.

## OUR OWN INVESTIGATIONS

When it is desired to investigate the stability of an intelligence test, two problems arise, *viz.*, (1) what correspondence or deviations there are between the results obtained when using the *same* test with varying temporal intervals between the tests, and (2) what correspondence or deviations there are between the results obtained when employing *different* tests with varying temporal intervals between the tests. In what follows we intend to present material for the elucidation of both these problems, and with one exception, we shall be dealing with adult subjects of 20 years of age.

## Re-test investigations with identical test

When determining the "accuracy" of measurement or the reliability of a test, as a rule, the split-half method is employed. It is relatively simple and requires little labour. The test is divided into two equivalent halves. Thus, to use Cronbach's (1949) terminology, a coefficient of equivalence is obtained, i.e., the correlation that would have been attained, had the subjects been simultaneously tested with two parallel scales. Any existing lack of correspondence is, with the exception of specificities, then due to factors which are active in the test situation itself; such as accidental success or failure, fluctuations in attention, misunderstanding, guesses, errors in registration and correction, etc. <sup>7)</sup>. In the split-half coefficient, however, factors that change from one test situation to another, such as general health, tiredness, motivation, the experimenter's personality, the character of the premises where the test takes place do not, however, affect the split-half coefficient. The influence that these factors exert depends on the degree of "constancy" of the subject <sup>8)</sup>. Since tests as a rule are meant to be predictive, it is of the greatest interest to determine the deviations that occur when the same subjects are tested with the same test with varying intervals between the administrations. In view of the purely administrative difficulties in obtaining sufficiently large and

---

<sup>7)</sup> For the sources of error that influence the split-half coefficient, vide Ekman (1947), p. 147 ff. and Husén (1949 a) p. 32 ff.

<sup>8)</sup> Cf. Ekman (1947 b) p. 144 ff.

representative groups of subjects at different times, the re-test method has been used to a much smaller extent than the split-half method for determining the reliability of a test. But as has been pointed out above, the split-half reliability may be quite high at the same time as the re-test reliability is very low. This is usually the case for individually administered tests for pre-school age. When tests are employed for predictive purposes, the determination of the re-test coefficients should be included as a standard part of the investigation of the usefulness of the tests in question.

Thus, the re-test coefficients obtained, vary with reference to the length of the re-test intervals and the subjects' ages. The longer the interval and the younger the subjects, the lower will be the re-test coefficient. R. L. Thorndike has made a survey of a large number of re-test investigations and plotted a curve for the observed values, which could be verified by subsequent investigations<sup>9)</sup>.

We give some examples. McMeeken (1939) carried out a re-test of 140 children of 10—11 years of age, with the Stanford-Binet Revision 1916, and with an interval of one day. The re-test coefficient obtained was  $+ .98$ , the highest coefficient that we have been able to find in the literature on this topic. The mean IQ difference was between 2.5 and 3 units<sup>10)</sup>. As a rule, the coefficients obtained for children are lower; and in this connection it should be noted, that generally the intervals between the tests were considerably longer than the one used by McMeeken. On the other hand, quite high re-test coefficients have been able to find in the literature on this topic. The Classification Test (AGCT), which was applied to 8,300,000 subjects during the period 1940—1944, shows a re-test coefficient of  $+ .90$  ( $n = 593$ ) with an interval varying from 1 to 13 weeks<sup>11)</sup>. Re-test coefficients for longer intervals are not given for this test. For 501 subjects who took the AGCT twice, at varying intervals, the correlation between the two series of test scores was  $+ .82$ . In the latter case, it was a question of a rather

---

<sup>9)</sup> R. L. Thorndike (1947). Also Ekman (1947 b) refers to a large number of retest investigations.

<sup>10)</sup> McMeeken (1939), p. 70.

<sup>11)</sup> Staff, Personnel Research Section, Psych. Bull. (42) 1945 p. 765.

selected group with only half the standard deviation of a normal population. If the standard errors of measurement that correspond to the 2 coefficients are worked out, it will be found that they are 7.7 and 5.1 units respectively. In other words: on account of the lower S.D. in the second case, the standard error of measurement is less, in spite of the considerably lower re-test coefficient. This illustrates another factor besides those mentioned above, and which ought to be taken into account when determining the reliability coefficients obtained, viz., the distribution of the test scores for the population investigated.

#### RE-TEST INVESTIGATION WITH I-TEST 1949.

A re-test investigation that was carried out under sufficiently controlled conditions, was based upon the *Induction Test of the Swedish Armed Forces*, the I-test 1949. This test has been described elsewhere, and consequently we have here only indicated its more general character <sup>12)</sup>. The test is composed of 4 segregated subtests. The split-half reliability is +.97. The number of items is 155, which are distributed among the four subtests as follows:

A. Directions	40
B. Synonyms	40
C. Concept Discrimination	40
D. Matrix	35

---

Total 155

The four subtests are of the following nature:

A. *Instructions*. The subjects are to carry out a task in accordance with special instructions. For example, "Draw a ring round the third letter of the fourth word in this sentence".

B. *Synonyms*. For a given word, the subject has to choose a synonym from five alternatives.

Ex.:

HIDE — lay out, elude, forget, conceal, decrease

---

<sup>12)</sup> Husén & Henricson (1951). „Directions” has been substituted for “Number Series”.



C. *Concept Discrimination*. From 5 given words, representing the same number of concepts, the subject has to indicate which word differs essentially from the other four.

Ex.: Water butter cream beer milk

D. *Matrix*. A portion of a larger figure has been "cut out". From 6 alternative pieces that are cut out, the subject has to choose the one that fits the "pattern" of the larger figure. This is a Swedish version of Raven's Matrix Test. The subjects of the re-test investigation were 136 recruits from the Södermanland mechanized regiment. In February-March 1949 they had been tested by 8 different local enlistment boards with a different tester at each board. When they were called up in June 1950, the subjects were tested for a second time with the same I-test and by one tester.

TABLE 1

Scatter diagram showing the retest correlation for I-test 1949. The first test was taken in February-March 1949, the second in June 1950. (Raw scores).

I-test 1949 (June 1950)

	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129	130-139	140-149	Total
0-9																
10-19																
20-29				1		1		1								3
30-39					1	1										2
40-49						3	2	3								8
50-59						1	5	5	1							12
60-69							3	5	1	2	1					12
70-79								1	2	3						6
80-89									5	6	4	1				16
90-99										1	5	3	1			10
100-109											3	4	1	1		9
110-119											1	2	12	4		19
120-129												1	5	13	4	23
130-139													2	4	8	14
140-149													1		1	2
Total:	—	—	—	1	1	6	10	15	9	12	14	11	22	22	13	136

$$r = +.95 \pm .01$$

The re-test correlation obtained for the total I-test is shown in Table 1. This coefficient is particularly high ( $+.946 \pm .009$ )

and is only two one-hundredth parts below the split-half coefficient, in spite of the fact that there was an interval of 15 months between the two test administrations. We had previously had reason to ask ourselves whether the split-half coefficient obtained might not be considered to be somewhat too high, as the test was time-limited, even if the limitation was rather moderate. The fact that the re-test coefficient is only slightly lower than the split-half coefficient, indicates that the latter is a correct expression for parallel test reliability.

As will be seen from the scatter diagram, the group is a positively selected sample. On induction, the mean raw score was 94.3 and S.D. was 31.2 as compared with 78.4 and 30.9 respectively for the age group as a whole<sup>13</sup>). The mean differs significantly from the national mean; while the S.D. does not diverge appreciably from that for the whole country. The fact that the material is selected to some extent, may have contributed towards producing a slightly higher re-test coefficient than would have been the case if the lower half of the variation range had been greater. It is found, and this is also apparent in Table 1, that the deviations tend to decrease as the ability level increases. Since subjects above the average are overrepresented, the re-test coefficient obtained may be expected to be rather on the large size. But in any case a decrease of more than 1—2 hundredths could hardly have been expected, even if the sample had been representative, for, as will be shown later, the correlation between two different group-tests, applied after an interval of one year, is over +.90.

If the coefficient +.95 is converted into a standard error of measurement by means of the formula  $\epsilon(a) = \sigma\sqrt{1-r}$ , we find that it is 7.1 raw score units or 3.5 IQ units.

There is a considerable difference between the test means and the S.D. of the first and second test administrations:

	M	S.D.
Feb.—March 1949	94.3 (IQ = 107)	31.2
June 1950	105.6 (IQ = 111.5)	28.2

The increase of the mean cannot very well be due to the

<sup>13</sup>) Husén: (1950 a).

subject being older at the second administration. In the spring of 1949, men born in 1929 and who were twenty years of age, were enlisted, and in the autumn of the same year, those born in 1930 were also enlisted. The mean difference in the scores was 0.5 in favour of the nineteen-year-old group. This difference was not significant<sup>14</sup>). The reason for the average increase was probably due both to the fact that on the second occasion the test was not new although the actual items had been forgotten, and that the conditions were also more favourable, inasmuch as the second test was carried out some time after the recruits had started their basic training, when they had had time to become more accustomed to their new environment. As a rule, an increase in IQ may be expected when a re-test takes place after an interval of one year or less. As has been stated above, McMeeken found that there was an increase of 2.5—3 IQ units when re-testing after one day with Stanford-Binet.

It is also of some interest to attempt to determine the re-test coefficients for the four subtests which make up the I-test. These coefficients are shown in Table 2 and are compared with the split-half coefficients.

TABLE 2

A comparison between the retest and the split-half coefficients of the subtests in the I-test 1949

Subtest:	Split-half coefficient <i>a</i>	Retest coefficient <i>b</i>	Difference <i>a-b</i>
A. Directions . . . . .	+ .92	+ .84	+ 0.08
B. Synonyms . . . . .	+ .96	+ .92	+ 0.04
C. Concept Discrimination . .	+ .92	+ .81	+ 0.11
D. Matrix . . . . .	+ .94	+ .87	+ 0.07
The whole test . . . . .	+ .97	+ .95	+ 0.02

The mean difference between split-half and retest coefficients for the subtests is 0.075.

#### RE-TEST INVESTIGATION WITH I-TEST 1948

A comprehensive standardization with control groups was carried out in October 1948 at the Naval Recruit School in

<sup>14</sup>) Husén: (1950 a) p. 18.

Karlskrona. One of the groups comprising 303 recruits was, *inter alia*, tested with the I-test 1948. All subjects had been tested 7—8 months previously with the same test in connection with their enlistment. The test-retest correlation was  $+ .90 \pm .01$ . This coefficient is significantly lower than the one obtained subsequently with the I-test 1949 after an interval of 15 months. At the same time, it should be pointed out, that the group which was tested with the I-test 1948 was distinctly more homogeneous (the standard deviation for the IQ was 12.5 as compared with 15 for the entire age-group).

The I-test 1948 has been fully described in another connection, and consequently we shall here only give the main characteristics of the test <sup>15)</sup>. It consisted of 4 segregated subtests. Both oral and written instructions were given before each subtest; and a time-limit was set for the solution of the items. The subtests with the number of items they contained were:

A. Synonyms	40 items
B. Concept Discrimination	40 „
C. Number Series	40 „
D. Matrices	40 „
<hr/>	
Total	160 items

The principal alterations that were made in the I-test 1949, as compared with that for 1948, consisted in Instructions being substituted for Number series; the directions for the various subtests were only partly given orally, and finally, Matrices was reduced by 5 items.

The first testing took place, as has already been mentioned, in connection with enlistment, and with about a dozen different testers. The re-test was carried out 50 days after the subjects had started their basic military training, and by three different testers.

The subjects tended to be a slightly negatively selected sample and the group was significantly more homogeneous than the whole corresponding age group.

<sup>15)</sup> Husén & Henricson (1951).



	M (IQ)	S.D.
Feb.-March 1948	96.4	12.5
Oct. 1948	101.3	11.1

The mean increased by 4.4 IQ-units for the second testing as compared with the first. The corresponding value for the I-test 1949 was 4.5 units. If we examine the scatter diagram (Table 3) more closely, we shall find that the subjects who were below average, and average, tended to have higher scores in the second testing; while the subjects belonging to the two highest deciles retained their status. That explains why the S.D. was lower in the second test than in the first. The same was observed in the re-test with the I-test 1949.

TABLE 3

Scatter diagram showing the retest correlation for the I-test 1948. The first test was taken in February-March 1948, and the second in October 1948.

(IQ-scores)

I-test 1948 (Febr.-March) IQ

IQ	68-70	71-73	74-76	77-79	80-82	83-85	86-88	89-91	92-94	95-97	98-100	101-103	104-106	107-109	110-112	113-115	116-118	119-121	122-124	125-127	Total
74-76	1	1	2																		4
77-79	1	1		2	1		1														6
80-82			4	2	2																8
83-85	1				2	3															6
86-88		1	2		2	7	2	2	1		1										18
89-91			1	2	2	4	1	1		1		1									13
92-94			2	1	3	2	3	7	3	3	1				1						26
95-97			1			1	11	6	7	3											29
98-100					1	1	4	4	6	6	3	2	1								28
101-103						1	1	5	11	7	8	4				2					39
104-106									2	5	8	6	3	1							25
107-109											6	4	10	7	1	1					29
110-112											1	3	6	6	3	3	3				25
113-115												1	2	5	2	7	1				18
116-118												1	2	1		5	5	2	1		17
119-121											1								1		2
122-124																	1	1		1	4
125-127																			1	1	2
128-130																			1	2	3
131-133																				1	1
Total	3	3	12	7	13	19	23	25	30	25	29	22	24	20	7	18	10	3	4	6	303

$$r = +.90 \pm .01$$

As has already been pointed out, the re-test coefficient,  $+.90 \pm .01$ , for the I-test 1948 is significantly lower than the corresponding coefficient for the I-test 1949 ( $+.95 \pm .01$ ). The reason for this is, *inter alia*, the greater homogeneity of the group, which was tested with the I-test 1948. By employing Kelley's formula for the correction of "restriction of range" we can calculate that a normal population with an IQ S.D. of 15, should have obtained a retest coefficient of about  $+.93$ <sup>16</sup>). A coefficient of  $+.90$  corresponds to a standard error of measurement of 4.0 IQ-units for the 1948 I-test. The corresponding value of the 1949 I-test was 3.5 units. Thus the difference in the standard error is quite small, in spite of the difference in the re-test coefficient.

On the basis of the re-test investigations described here, it seems reasonable to require that a group-test scale for adults should, after an interval of about 1 year, give for a normal population a re-test coefficient exceeding  $+.90$ , or to express this differently, the standard error of measurement in IQ-units should not be greater than 4.7. It may be pointed out that in a large number of re-test investigations, where the best standardized individual test in existence, Stanford-Binet 1937, was used, the standard error of measurement was observed to be approximately 5 IQ-units, even when the intervals were of relatively moderate length<sup>17</sup>).

#### Re-test investigations with the application of different tests

From the operational point of view, which has for a long time been self-evident in physics, but has only of late begun to penetrate psychological thinking, it is clear that no general trait exists that may be labelled "intelligence". On the contrary, there are several kinds of intelligence, each of which is defined in accordance with the methods (tests, marks) that constitute the basis for observation. There is, of course, nothing to prevent us from defining the quality that is measured by a certain test as

---

<sup>16</sup>) Guilford (1936), p. 416.

<sup>17</sup>) Ekman (1947 b), p. 79 ff.

intelligence. In the same way, for example, in Sweden we call what we measure with a centigrade thermometer, the temperature. At present, however, no such generally accepted "scale of reference" exists for measuring intelligence<sup>18)</sup>. A large number of both individual and group tests are in use, each of which measures a certain kind of intelligence which overlaps more or less with intelligence as measured by other tests. Therefore it is naturally of the greatest importance to determine to what extent these different scales measure the same thing; in order to be able, among other things, to convert the scores obtained by one test into the probable scores of another test. In view of the great importance attached to the intelligence test results, e.g. in connection with legal procedure and social welfare, a thorough comparative investigation would prove of very considerable value. As regards the purely technical side, it should be arranged in such a way that a normal population would have to be tested at relatively the same time, with a number of different intelligence tests. So far, such an investigation has not been undertaken. On the other hand, there have been several investigations on a smaller scale, where the results of various Swedish intelligence tests have been correlated with one another, as a rule, two at a time.

It is of interest, however, to correlate the scores not only of tests applied simultaneously or nearly simultaneously, but also to correlate the results of different tests applied both after longer and shorter intervals. As has been emphasized above, intelligence tests are, as a rule, used to predict future achievements or social adaptability. Just as it is not sufficient to determine the split-half reliability of a test, i.e., its momentary stability, without taking into account its "stability" over a long period, it is not enough to determine the correlation of two tests simultaneously administrations. It happens, for example, in forensic psychiatry this correlation is affected by an interval occurring between the administrations. It happens for example, in forensic psychiatry in Sweden, that patients are tested by means of different tests with longer or shorter intervals. In such a case, it is of decisive importance to know the distribution of the deviations<sup>19)</sup>.

<sup>18)</sup> It has been urged from various quarters that by IQ should be understood the intelligence determined by means of the Terman-Merrill test.

<sup>19)</sup> Cf. Ekman (1950)

The investigations described below are mainly concerned with adult subjects. These investigations have, with few exceptions, been arranged with the definite purpose of determining the correlations between two or more of the intelligence tests in question. We have found it expedient, however, to compare those results, which since in many cases they refer to normal populations, will serve well to throw light on the stability of the tests.

The following intelligence tests have been used:

Anderberg's group-test scale 1947

I-test 1944, 1945, 1946 and 1948.

CVB-Scale 1946 and 1949.

Terman-Merrill 1943 (Swedish version)

Point-Scale (Lindberg).

TABLE 4  
Certain numerical data relating to the various tests.

Test:	General characteristics:	Number of items:	Time allowance (including instructions):	Reliability (Sp = split-half (Rt = retest)
Anderberg's group test 1947	Omnibus-scale, divided into two halves, many different types of items <sup>1)</sup>	49	unlimited	+ .90 (Rt)
I-test 1944	Segregated, 8 subtests <sup>2)</sup>	157	47 min	+ .95 (Sp)
I-test 1945 group	" " " <sup>3)</sup>	"	47 min	+ .96 (Sp)
I-test 1946 tests	" " " <sup>3)</sup>	"	47 min	+ .96 (Sp)
I-test 1948	" 4 subtests <sup>4)</sup>	160	61 min	+ .97 (Sp)
CVB-Scale 1946	Individual test			
	8 subtests, segregated <sup>5)</sup>	119	about 52 min	+ .95 (Sp)
CVB-Scale 1949	" <sup>6)</sup>	108	" 50 min	.96 (Sp)
CVB-Scale half-scale I	" <sup>7)</sup>	54	" 28 min	+ .93 (Sp)
Terman-Merrill (Swedish version)	Individual test graded according to age, 6 items for each age level <sup>8)</sup>	—	" 50 min	+ .93 (Rt)
Point-Scale (Lindberg)	Individual test segregated, 9 groups of items <sup>9)</sup>	50	" 25 min	+ .88 (Rt)

<sup>1)</sup> Anderberg (1947) and Agrell (1950) p. 48

<sup>2)</sup> Husén & Ekman (1944), p. 118 ff.

<sup>3)</sup> Husén (1946 a)

<sup>4)</sup> Husén & Henriksen (1951)

<sup>5)</sup> Husén (1946 b)

<sup>6)</sup> Husén (1949 b)

<sup>7)</sup> Husén (1949 b)

<sup>8)</sup> Terman-Merrill (1937)

<sup>9)</sup> Lindberg (1943), p. 5



These are the tests most commonly used in Sweden for adults. This further stresses the importance of investigating the magnitude of the overlapping. Before we turn to the actual investigations, it may be useful to tabulate (Table 4) the characteristics of the tests. As regards the type of the items, the reader is referred to the publications in question.

As has already been mentioned, we are going to compare the tests two at a time. Only in one case have the same subjects been tested by means of three tests. We have thus, not compared each test with all the others, but have restricted the number of

TABLE 5

Investigations of the testing of the same subjects by means of two different tests, and where the tests have been correlated with each other.

Tests compared:	Subjects:	Intervals between test administrations:	Correlations between the scores:
CVB-Scale 1946-Terman-Merrill	53 patients at the department for forensic psychiatry at Ulleråker (Mean IQ = 89)	about 1 week	+ .91 ± .02
CVB-Scale 1946-Terman-Merrill	89 patients at the department for forensic psychiatry at Lillhagen (Mean IQ = 89)	„ „	+ .94 ± .01
CVB-Scale 1946-Point-Scale	62 patients at the department for forensic psychiatry at Ulleråker (Mean IQ = 88)	„ „	+ .80 ± .04
CVB-Scale 1946-I-test 1944	100 recruits aged 21; controlled, representative sample	18 months	+ .83 ± .03
CVB-Scale 1949-I-test 1947	180 recruits aged 21; controlled, representative sample	14-22 „	+ .84 ± .03
CVB-Scale 1949-half scale I-I-test 1948	101 recruits aged 21, sentenced, slightly negatively selected sample (Mean IQ = 97)	19-20 „	+ .87 ± .02
Anderberg's group test 1947-I-test 1945	103 recruits aged 21; positively selected sample (Mean IQ = 108)	24 „	+ .83 ± .03
I-test 1946-I-test 1948	442 recruits aged 21; controlled, representative sample	14 „	+ .91 ± .01
Hallgren's group test-I-test 1948	613 recruits, tested in 3rd class of elementary school and at enlistment	10 years	+ .72 ± .02

comparisons to 9. Before giving a detailed account of the investigations where possibilities for comparison existed, we shall start by tabulating the investigations (Table 5).

The number of comparisons made is 9, 3 of which refer to two individual tests, 3 refer to one individual test and a group, and 3 to two group tests. If we exclude the comparisons with the Point Scale, since the distribution for this test is extremely skewed, we find that the correlations within the group tests and the individual tests respectively tend to exceed  $+ .90$ , while the correlation between an individual and a group test tends to be between  $+ .80$  and  $+ .90$ . The various correlation coefficients are not strictly comparable for the following reasons:

- (1) The intervals between the two test administrations vary considerably.
- (2) The samples are very varied as regards their composition. The mean, S.D. and range of variation, differ very considerably in the various groups.
- (3) The conditions under which the different tests were given varied. Some tests were taken in connection with the enlistment of recruits, others were taken in connection with forensic psychiatry or with military service.

In order to assess the significance of the various coefficients properly, it is necessary to specify in greater detail the conditions outlined under the three headings above. Where it was possible to determine the coefficients for the various subtests, these data will also be given in what follows.

A. *CVB-Scale 1946 — Terman-Merrill.* ( $+ .91$ ) Here the group consisted of 53 patients from the department for forensic psychiatry of the Ulleråker hospital. Both variables are available as IQ-scores. Both tests were taken during the course of the same psychiatric examination period. The mean interval was estimated to be about 1 week. The means and standard deviations were

	M	S.D.
CVB 1946 (IQ)	89.8	12.4
Terman-Merrill (IQ)	89.2	19.0

As will be seen the S.D. for the Terman-Merrill test is considerably greater than that of the CVB-Scale. That is the

explanation why the values of the former are too high or too low at both extremes of the distribution. It shows that the Swedish version of the Terman-Merrill test does not meet the demands of its originator, that the S.D. should equal 17.0. Consequently the Swedish version of the Terman-Merrill test gives IQ-scores whose significance is different from that of IQ's obtained by other tests.

B. *CVB-Scale 1946 — Terman-Merrill (+.94)* The group consisted of 89 patients at the Lillhagen Hospital (department of forensic psychiatry). Both variables were available as IQ's. Both tests were taken in the course of one and the same psychiatric examination.

The means and S.D.'s for the two tests are as follows:

	M	S.D.
CVB (IQ)	96.3	14.9
Terman-Merrill	95.4	22.9

As in comparison A, the S.D. for the Terman-Merrill test is considerably greater than that for the CVB-Scale. On the basis of the data given here, we can determine the two regression equations which are required for estimating the probable values of the one test that correspond with the given values of another. The two equations are:

$$Y = 0.61x + 23.7;$$

$$X = 1.44y - 23.3.$$

The two regression lines intersect when the IQ equals about 80. The size of the deviations between the two tests is such as to be of some practical significance for scores for either test above 70 or below 110. The significance of the greater S.D. of the IQ of the Terman-Merrill test has been considered above.

C. *CVB-Scale 1946 — Point-Scale (+.80)* The group consisted of 62 patients from the department of forensic psychiatry at Ulleråker hospital. The values of the CVB-Scale are available as IQ, and those of the Point-Scale as raw scores. Both tests were taken during one and the same psychiatric examination period. The means and standard deviations were:

	M	S.D.
CVB 1946 (IQ)	87.7	12.4
Point-Scale (raw scores)	37.7	7.6

The value 37.7 for the Point-Scale corresponds, according to Lindberg (1943), to an M.A. = 13 years. An important reason for the relatively low correlation between the two tests is the particularly skewed distribution of the Point-Scale scores <sup>20</sup>). Although this was a negatively selected group, 45 subjects out of 62 are bracketed within the interval 40 to 50 points. There is thus no question of any differentiation upwards in this test.

D. *CVB-Scale 1946 — I-test 1944. (+.83)* The group consists of 62 are bracketed within the interval 40 to 50 points. There recruits born in 1924. The first test (I-test 1944) was taken in 15 different enlistment areas in February—March 1944. The second test (CVB-Scale) was taken in October 1945 with 4 different examiners <sup>21</sup>).

The means and S.D.'s were as follows

	M	S.D.
CVB 1946 (IQ)	100.2	15.1
I-test 1944 (raw scores)	70.9	33.1

E. *CVB-Scale 1949 — I-test 1947 (+.85)* The group consisted of a controlled, representative sample of 180 conscripts belonging to the 1947 class, i.e. born in 1927. The first test (I-test 1947) was taken in connection with their enlistment in February-March 1947 in about 10 enlistment areas. The second test (CVB-Scale) was carried out in May—June and November—December 1948 and two different testers administered the test <sup>22</sup>). The means and standard deviations of the two tests were:

	M	S.D.
CVB (IQ)	100.0	15.1
I-test 1947 (raw scores)	81.8	32.6

<sup>20</sup>) On account of the skewed distribution of Point-Scale scores it is not very appropriate to use product moment correlation coefficients.

<sup>21</sup>) Husén (1946 b) p. 18.

<sup>22</sup>) Husén (1949 b).



TABLE 6

Scatter diagram showing the correlation between the I-test 1947 and the CVB-Scale 1949; in both cases raw scores are given. The first test was taken in February-March 1947, and the second in November 1948.

I-test 1947 (raw scores)

	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129	130-139	140-149	Total
CVB-Scale (raw scores)																
20-29	1			1		1										3
30-39	1															1
40-49			1		1											2
50-59	1		1		1											3
60-69		1	1													2
70-79		1	1			1										3
80-89		2	2			1	1									6
90-99				3	3	1	2	1								10
100-109				2		2	3	3	1	2						13
110-119					1	1	3	4	2			1				12
120-129				1	1	1	6	3	4	2						18
130-139				1	2	2	2	4	9		2					22
140-149						1		5	2	5	5	1	1			20
150-159								2	3	2	2	1				10
160-169										5	4	3	2	1		15
170-179						1		1	1	1	2	4	1	1		12
180-189										2	3	3	1	1		10
190-199											1	1	3	3		8
200-209												1	2	3	1	7
210-219													1		2	3
Total	3	4	6	8	9	12	17	23	22	19	19	15	11	9	3	180

$$r = +.85 \pm .02$$

F. CVB-Scale 1949 (part I) — I-test 1948. (+.87). The group consists of 101 recruits undergoing disciplinary punishment from 9 different units born 1928. The first test was taken in Feb.—March 1948 at 15 different enlistment areas using I-test 1948. IQ-scores are available. The test has a split-half-reliability of +.97. The second testing (CVB-Scale, part I) took place in October—November 1949 in connection with an individual examination of the recruits having undergone disciplinary punishment. The CVB-Scale consists of two equivalent parts, each with a split-half-reliability of +.93. In the individual

testing there were only two testers. The means and S.D.'s of the two tests were:

	M	S.D.
CVB 1949, part I (IQ)	96.2	14.3
I-test 1948 (IQ)	96.7	13.6

It is possible, by means of the formula for the correction of attenuation, to determine that the correlation between them should have been  $+ .91$  as against  $+ .87$  for the non-corrected coefficient, had the two tests been perfectly reliable.

TABLE 7

Scatter diagram showing the correlation between IQ according to the I-test 1948, and IQ according to part I of the CVB-Scale. The first test was taken in February–March 1948, and the second in November 1949.

The half-scale of the CVB-Scale (IQ)

	65–69	70–74	75–79	80–84	85–89	90–94	95–99	100–104	105–109	110–114	115–119	120–124	125–129	130–134	135–139	Total
I-test 1948 (IQ)																
65–69			1													1
70–74	1		2		1											4
75–79		1	1	2	1											5
80–84		1	3	4	1	1										10
85–89			2	5	2	1	1		1							12
90–94				3	2	2	4	2								13
95–99					3	6	2	1	2							14
100–104					1	3	3	2	4	1		1				15
105–109							2	1	2	4	1					10
110–114								1			1	2				5
115–119										3	1	1				5
120–124									1	1	1	2				5
125–129											1			1		2
130–134																
Total	1	2	9	14	11	13	12	7	11	9	5	6	—	1	—	101

$$r = + .87 \pm .02$$

G. *Anderberg's Group test 1947 — I-test 1945.* ( $+ .83$ ) The group consisted of 103 recruits of the 8th Infantry Regiment (I 8) born in 1925. They were tested at their enlistment in February–March 1945; on this occasion practically all of them were tested by the same tester. The split-half reliability for the I-test 1945 is  $+ .96$ . Anderberg's group test, which was given to

the recruits in March 1947, has been described by Agrell (1950), who also conducted the comparison referred to here. The correlation between the two tests is lower than what is usually found between group-test scales. The reasons for this were, among other things, as follows:

(1) The interval between the two test administrations was more than two years, which is longer than that of any of the other comparisons so far treated here.

(2) The group was positively selected (the mean raw score was 100.5 as against 80.6 for the entire age group, and was also more homogeneous than the age group as a whole  $S.D. = 27.1$  as compared with 31.6). The variation area 0—30 raw scores was thus not at all represented (cf. Agrell, 1950 pp. 48—49).

H. *I-test 1946 — I-test 1948.* (+.91) The group consisted of 442 recruits from 5 different units, born in 1926. They were tested in February—March 1946 at 17 different enlistment centres. The second test was administered in May—June 1947 in connection with the standardization of the I-test 1948, and three different testers were employed. The split-half reliability for the I-test 1946 is +.96 and for the I-test 1948 +.97. The means and standard deviations for the two tests are:

	M	S.D.
I-test 1946 (raw scores)	80.8	30.9
I-test 1948 (raw scores)	77.0	31.8

Thus, as regards the results of the I-test, the group is on the whole representative, since the mean for the whole age group born 1926 was 82.6 and the S.D. was 31.4. The mean for the whole age class 1948 was 70.5 and the S.D. was 34.1. In our material the corresponding values were (see above) 77.0 and 31.8 respectively. But here we have to take into account the *effect of training*, due to the I-test that had been administered previously. In a retest with an *identical* test we have found that there is an average increase of 4—4.5 IQ units or 9—11 raw scores, and that at the same time the S.D. is lowered by about 3 raw score units. The tested group was, in the I-test 1946, 1.8 points below the level of the corresponding entire age group. This value plus the difference between 77.0 and 70.5 amounts to 8.3 points.

TABLE 8

Scatter diagram showing the correlation between the I-test 1946 and the I-test 1948. The first test was taken in February-March 1946, the second in May 1947. (Raw scores)

I-test 1946 (raw scores)

	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129	130-139	140-149	Total
I-test 1948 (raw scores)																
0-9		2	1													3
10-19	3	4	2						1							10
20-29	2		6	2	1	2	1									14
30-39		1	7	2	8	9	3	2			2					34
40-49			1	5	10	9	4	4	1		1					35
50-59			1	4	6	11	11	6	8		1					48
60-69					1	8	15	11	4	4						43
70-79					2	1	11	11	11	3	1		1			41
80-89						2	2	7	13	11	11	1				47
90-99								3	6	20	15	9				53
100-109								1	5	9	19	10				44
110-119										3	6	8	7			24
120-129											3	8	4	5	1	21
130-139												1	5	11	2	19
140-149														4	2	6
Total	5	7	18	13	28	42	47	45	49	50	59	37	17	20	5	442

$$r = +.91 \pm .01$$

The difference in S.D. in the I-test 1948, between our group and the entire age group, is 2.3 points, i.e. a decrease that is slightly less than that obtained when retesting with an identical test. In terms of the I-test 1946, the group was somewhat more homogeneous from the outset than the entire age group. (In the comparisons referred to here, no account has been taken of sampling errors.)

By way of summary, it may be stated:

1. When retesting representative groups from the normal population of adult subjects by means of *identical* group tests, and at an interval of about 1 year, we have the right to expect or even to demand retest coefficients that exceed  $+.90$ .
2. When retesting representative groups of adult subjects by means of *different* tests, and at an interval of about 1 year, we could expect to obtain coefficients at the  $+.85$  level.



- a) If two group tests are compared, a correlation of about  $+.90$  may be expected.
- b) If two individual tests are compared, a correlation of about  $+.90$  may also be expected.
- c) If an individual test and a group test are compared, a correlation of about  $+.85$  or over, may be expected.

3. The Swedish version of Terman-Merrill shows for adults a S.D. of over 20, which implies that IQ values, especially at the extremes of the distribution of scores, have a different significance from those in the other IQ-standard scales, namely, the values obtained at the negative side are too low and the values at the positive side are too high.

### ZUSAMMENFASSUNG

#### *Ueber die Stabilität der Intelligenztestwerte*

1. Der Verf. weist darauf hin, dass die „Zuverlässigkeit“ einer Testmethode drei verschiedene Aspekte bietet. Erstens handelt es sich um die sog. Validität, das heisst in welchem Umfange der Test dasselbe wie das Kriterium misst. Zweitens haben wir mit der Präzision (Reliabilität) zu tun, das heisst wieviel zufällige Umständen die Werte beeinflussen. Drittens handelt es sich um die „Stabilität“ der Testwerte, das heisst wie gross die Korrelation zwischen zwei Serien von Messungen von derselben Vpn., die mit demselben oder zwei verschiedenen Testinstrumenten geprüft werden, ist. Bei der Konstruktion von Testreihen werden die beiden ersten Probleme in der Regel bemerkt. Viel seltener beachtet man das Problem der Stabilität. Man ist aber oft in der Lage, dass die Stabilität untersucht werden muss, z.B. wenn man den prognostischen Wert einer Testreihe für Kinder berechnen will. Es kommt oft vor, dass die Präzision ziemlich gross ist, was auch auf kurze Sicht mit der Validität der Fall sein kann. Die Stabilität kann aber sehr niedrig sein, was die Validität auf lange Sicht erheblich reduzieren kann.

2. Die bisherigen Untersuchungen sind diskutiert und es wird die Schlussfolgerung gezogen, dass zwei Faktoren die Stabilität beeinflussen, erstens die Länge des Zeitraumes zwischen den beiden Messungen und zweitens das Alter der Vpn, bei der

ersten Messung. In Anschluss an J. E. Anderson weist der Verf. darauf hin, dass die Stabilität um so grösser wird, je älter die Vpn. bei der ersten Messung sind. Das hängt mit einfachen statistischen Tatsachen zusammen. (Die Menge von gemeinsamen Elementen.)

3. Eine Reihe von Experimenten, die mit Erwachsenen ausgeführt sind, werden beschrieben. Es handelt sich in sämtlichen Fällen um zwei Messungen von derselben Population mit verschiedenen Zeitintervallen. In zwei Untersuchungen sind die Vpn. mit *derselben* Testreihen und nach  $1\frac{1}{3}$  Jahr ein zweites Mal geprüft. In neun Fällen handelt es sich um Untersuchungen mit zwei *verschiedenen* Testreihen in jedem Paar von Werten, die verglichen werden. Die Testreihen sind in Tab. 4 beschrieben und alle sind in Schweden in praktischem Gebrauch.

4. Mit der Einschreibungsprobe (Version 1949) der schwedischen Wehrmacht wurden 136 Vpn. zweimal mit einem Intervall von 16 Monaten geprüft. Die Korrelation zwischen den beiden Serien Testwerte war  $+0.95 \pm 0.01$ . Die Reliabilität (mittels der Hälftemethode) war 0.97. In einer zweiten Untersuchung wurden 303 Vpn. zweimal mit einem Intervall von 8 Monaten mit der Einschreibungsprobe Version 1948 geprüft. Die Korrelation erreichte diesmal die Höhe von  $+0.90 \pm 0.01$ . Es zeigt sich also, dass man mit diesem Typus von Gruppentests eine verhältnismässig hohe Stabilität erwarten kann.

5. In Schweden, wie in anderen Ländern, werden eine Reihe von Individualtests und Gruppentests für praktische Zwecke benutzt. Diese Proben definieren jede für sich eine Form von „Intelligenz“. Es ist in diesem Zusammenhang von grossem Interesse festzustellen, nicht nur wie gross die Stabilität, sondern auch wie gross die Ueberdeckung („overlapping“) zwischen den verschiedenen Proben sein mag. Wenn man den Umfang der gemeinsamen Elemente feststellt kann man die Werte der einen Probe in Werten einer anderen Probe umwandeln. Folgende Testreihen wurden dabei benutzt und folgende Vergleichen wurden gemacht:

a. Anderbergs Gruppentest wurde mit der Einschreibungsprobe 1945 verglichen. 103 Rekruten wurden mit einem Intervall von 24 Monaten geprüft. Die Korrelation zwischen den beiden Testreihen war  $+0.83$ . Der Koeffizient ist wegen geringer Streuung der Population ziemlich klein.

b. Die CVB-Skala, eine Individualtestreihe für Erwachsene, wurde in zwei Gruppen von Vpn. an einer psychiatrischen Klinik mit Terman-Merrill's Individualtestreihe verglichen. Das Intervall war eine Woche. Die Korrelationen wurden  $+0.94$  bzw.  $+0.91$ .

c. Die Einschreibungsprobe 1946 wurde in einer Gruppe von 442 Vpn. (Rekruten) mit Einschreibungsprobe 1948 verglichen. Das Intervall war 14 Monate. Die Korrelation war 0.91.

d. In vier andren Vergleichen zwischen die CVB-Skala (in ganzem oder die eine Hälfte davon) erhielten wir nach einem Intervall von 14—22 Monaten Korrelationen in der Höhe von  $+0.83$  bis  $+0.87$  mit verschiedenen Versionen der Einschreibungsprobe.

6. Von diesen Untersuchungen werden diese Schlussfolgerungen gezogen:

a. Wenn eine verhältnismässig repräsentative Population mit einem Intervall von einigen Monaten zweimal mit einer *identischen* Gruppentest konventioneller Art geprüft wird, kann man eine Korrelation zwischen den beiden Serien Testwerte von mindestens  $+0.90$  erwarten.

b. Wenn eine Population mit einem Intervall von einigen Monaten zweimal mit zwei *verschiedenen* Tests geprüft wird, kann man eine Korrelation von etwa  $+0.85$  erwarten. Sind die Testreihen beide Gruppentests (oder Individualtests), kann man eine Korrelation von etwa  $+0.90$  erwarten.

7. Die vorliegende Untersuchung hat gezeigt, dass die Testreihen, die der elementaren Erfordernisse hinsichtlich üblichen Konstruktionsprinzipien entsprechen, in ziemlich grossen Umfange dieselben Dinge messen. Die Ueberdeckung, die wir feststellen können, hängt gewiss in ziemlich grossem Masse damit zusammen, dass die Validität der meisten Intelligenztestreihen mit Schulleistungen und gleichen Prestationen festgestellt wird.

#### REFERENCES

- Agrell, J. (1950), „Skolreformen och näringslivet“. Stockholm (in Swedish).  
Anastasi, A. & Foley, J. P. (1949), „Differential psychology“. New York.  
Anderberg, R. (1947), „Grupptest för intelligensundersökning (15 år och högre åldrar)“. Uppsala (in Swedish).  
Anderson, J. E. (1940), The predication of terminal intelligence from infant and preschool tests. *Yearbook Nat. Soc. Stud. Educ.* (39:I).

- Anderson, L. D. (1939), The predictive value of infancy tests in relation to intelligence at five years. *Child Development* (10).
- Bayley, N. (1940), Mental growth in young children. *Yearbook Nat. Soc. Stud. Educ.* (39:II).
- (1948), Consistency and variability of mental test scores from three months through eighteen years for a constant case sample. *American Psychologist* (2).
- Brown, R. R. (1933), Time-interval and constancy of the IQ. *Journ. of Educ. Psych.* (24).
- Cronbach, L. J. (1949), „Essentials of psychological testing“. New York.
- Ekman, G. (1947 a), „Intelligensmätning“. Uppsala (in Swedish).
- (1947 b), „Reliabilitet och konstans“. Uppsala (in Swedish).
- (1950), Om intelligensmätningarnas innebörd. *Svenska Läkartidningen* No 19 (in Swedish).
- Elonen, A. (1949), A comparison of two tests of intelligence administered to adults. *Psychological Monographs* (63) No 11.
- Embree, R. B. (1948), The status of college students in terms of IQ's determined during childhood. *American Psychologist* (2).
- Freeman, F. S. (1950), „Theory and practice of psychological testing“. New York.
- Guilford, J. P. (1936), „Psychometric methods“. New York.
- Hildreth, G. H. (1926), Stanford-Binet retests of 441 school children. *Ped. Sem.* (33).
- Honzik, M. P. (1938), The constancy of mental test performance during the preschool period. *Journ. of Genet. Psych.* (52).
- Husén, T. (1946 a), „Några data rörande svenska krigsmaktens inskrivningsprov“ 1 : st Ed. (1946 a) 2 : nd Ed. (1947) 3 : rd Ed. (1948) 4 : th Ed. (1950 a) Lund, (in Swedish).
- (1946 b), „CVB-skalan. Individualtestskala för vuxna“. Lund (in Swedish).
- (1948), „Konstruktion och standardisering av svenska krigsmaktens inskrivningsprov“. Lund (in Swedish).
- (1949 a), „Om innebörden av psykologiska mätningar“. Lund (in Swedish).
- (1949 b), „CVB-Skalan. Individualtestskala för vuxna. Reviderad upplaga“. Lund (in Swedish).
- (1950), „Testresultatens prognosvärde“. Lund (in Swedish).
- (1951), The influence of schooling upon IQ. *Theoria* (17). Lund: Gleerup.
- & Ekman, G. (1944), Inskrivningsprovet 1944. *Tidskrift för psykologi och pedagogik*, (2). Gothenburg (in Swedish).
- & Henricson, S. E. (1951), „Some principles of construction of group intelligence tests for adults“. Stockholm: Almqvist & Wiksell.
- Jones, H. E. (1946), Environmental influences on mental development. In Carmichael: *Manual of child psychology*. New York.
- Lindberg, B. J. (1943), „Intelligensbestämning enligt „Point-Scale“ och Wåhléns metod“. Gothenburg (in Swedish).



McMeeken, A. M. (1939), „The intelligence of a representative group of Scottish children". London.

Terman, L. M. & Merrill, M. A. (1943), „Intelligensmätning". Stockholm (in Swedish).

The Army general classification test. *Psych. Bull.* (42) 1945.

Thorndike, R. L. (1933), The effect of the interval between test and retest on the constancy of the IQ. *Journ. of Educ. Psych.* (24).

——— (1940), „Constancy" of the IQ. *Psych. Bull.* (37).

——— (1947), The prediction of intelligence at college entrance from earlier tests. *Journ. of Educ. Psych.* (38).

——— (1950), Individual differences. In: *Annual Review of Psychology*. (1). Stanford.

## SEX DIFFERENCES IN LOCALIZATION AND ORIENTATION

BY

CARL IVAR SANDSTRÖM  
*University of Stockholm*

In an earlier work the author has described a phenomenon in the field of opto-kinesthetics, which had not been previously observed; in this work the author endeavoured to investigate more fully its nature and scope <sup>1)</sup>. Among the findings that resulted from the experimental work were certain sex differences that were connected with the phenomenon. Since sex differences in the psychology of perception are a matter of quite general interest, we consider it expedient to group together our results, particularly from this point of view, and to make them available in a shorter paper.

### 1

The phenomenon may be described briefly as follows: A *luminous point that does not give any guiding light in an otherwise completely dark room, and which is situated at a convenient distance for pointing (about 50 cm), cannot be correctly pointed to and touched (for instance with the forefinger) by a direct pointing movement.* The phenomenon is valid for both binocular and monocular sight. In the experiments described only ordinary binocular sight is taken into account.

To give a concrete example: in a dark room and at convenient distance for pointing, it is not possible to point to and touch a little hole that has been made by a drawing pin or some similar object in a board used for blackout purposes, and through which hole the daylight is shining clearly and distinctly. Practically without exception, this task is looked upon as being very easy to accomplish and it is a foregone conclusion that the result will

---

<sup>1)</sup> Sandström, C. I. *Orientation in the Present Space*. Stockholm: Almqvist & Wiksell 1951.

prove successful. The fact that the point continues to shine even after the pointing operation has been concluded is experienced by the S as quite astonishing. One of our first Ss gave the following account of his experience: "It appears that the luminous point shines right through one's finger". This aptly describes the assurance which is felt that one has touched the right spot; even though this first phase of the experience of failure only rarely finds spontaneous expression except in the case of phenomenologically well-trained observers. Our experience — for light cannot shine through a finger — appears immediately to correct our error, and this implies that we can indicate the position of our point of impact in relation to the luminous point. To this indication of location, which is often felt to be very definite, we shall have occasion to return in section 4.

It should be pointed out that the phenomenon is not connected with difficulties in estimating the distance. Immediately before carrying out the experiment, the S is free to choose, under normal lighting conditions, his own optimal distance for pointing and also to practise the movements required for this. Between each trial the light is switched on, and it is only switched off just before the experiment is conducted. No S has reported either spontaneously or in answer to a question, that he touched the target sooner or later than he had expected. It appears quite impossible to interpret the deviation errors as due to such misconceptions<sup>2)</sup>. A proof of this, is the "pointing-from-below" experiment, described below, which was carried out when the room was fully lit.

## 2

The apparatus employed was of a simple character.

A source of light consisting of a 15 Watts electric bulb was built into a wooden box  $30 \times 30 \times 10$  cm. A masonite board 60 cm square was attached to the front of the box, and functioned as a sliding lid. A small hole was made in the centre of the

---

<sup>2)</sup> We would remind the reader of Bourdon's classical experiments in connection with the estimation of distance in the dark. (Bourdon, B. *La perception visuelle de l'espace*. Paris 1902) Cf. also Woodworth, R.S. *Experimental Psychology*. New York: Henry Holt & Co 1938 p. 670.

masonite board. A fine metal tube, similar to those used for the tips of hypodermic syringes, was inserted in the hole. The diameter of the tube was 0.65 mm. What are known as (short) mapping pins were used for pointing, and to start with, the Ss were advised to stand in a relaxed position with their arms at their sides. The Ss had to stick the pin on the board at the first point of impact. Millimeter paper was attached to the board, and the position of the point of impact in relation to the target was determined by means of the values obtained on the abscissa and on the ordinate.

The experiments which are dealt with, formed part of two comprehensive experimental series in which different Ss were employed.

66 Ss (34 male and 32 female) took part in the first series, and 48 Ss (23 male and 25 female) in the second series. All the Ss were students of psychology. Every subtest was carried out four times during the session. In this the S was not shown his result, so that he was quite unable to make any corrections on the strength of the experience he had acquired.

Before considering the results, we shall first briefly describe some parallel experiments. Working with the same contrivance, and at the same distance, as were used for the luminous-point experiment, the S was required when employing binocular vision, and with the room normally lighted, to fix a small red fixation point, which was situated at the same place as the luminous point. The S was then asked to shut his eyes and with the aid of his spatial memory-image to attempt to stick a mapping pin into the fixation point<sup>3)</sup>. The experiment is referred to as *the pointing-with-shut-eyes experiment*.

In another experiment, which formed a part of the latter series, the S was sitting in front of a specially constructed table. The table-top, which consisted of a piece of plywood, was painted white, and on it a black fixation point was marked, and while fixing this point binocularly, the S had to stick a mapping pin on exactly the corresponding spot on the underside of the table-top. This is referred to as *the pointing-from-below experiment*.

---

<sup>3)</sup> It appears that Bowditch and Southard (J. Physiol. 3: 232—245, 1880) were the first to conduct experiments of this type. For a synopsis, see Lawlor, Arch. Psychol. No. 213 (1937).



The luminous-point experiment in the second series of experiments — besides pointing to a white light as had been done previously — also consisted in pointing to a red light. A frame was built into the box, so that a red colour filter could be inserted and also readily withdrawn when the white light was required. During a special session all the 48 Ss carried out the luminous-point experiment with both white and red light under conditions of dark-adaptation. The purpose of these variations was to ascertain whether the errors in pointing could be due to retinal changes caused by the darkness. This proved not to be the case; and the pointing-from-below experiment furnishes a proof that these errors are dependent on other factors; above all on the fact that the pointing instrument is not visible on account of the dark, also this absence of the possibility of visual control is not experienced as a deficiency <sup>4)</sup>.

---

<sup>4)</sup> With regard to the height at which the target was placed, the following may be stated, since also here was a difference due to sex, that we found surprising. At the outset it seemed obvious that the target should be placed at the objective eye-level; but during the preparatory experiments a number of the Ss stated — this was especially the case among the male Ss — that the target was too high and was experienced as being above eye-level. As we were desirous of having the conditions as natural as possible, we let the Ss themselves choose the „pointing-height” in accordance with the following method.

A strip of paper was fixed vertically to one of the walls of the experimental room. A row of small dots was marked on the paper, and the Ss while standing in front of the strip of paper at a distance of about 50 cm, were requested to point to and touch a point that seemed most „convenient” to them. The distance between the dots on the paper was 2 cm. With only one exception, none of the Ss found this test either peculiar or difficult to accomplish, but pointed to one of the dots with considerable assurance. In the above-mentioned first series of experiments the target was placed at the same height as the dot that had been selected by the Ss.

The objective eye-level of the Ss was measured and the average eye-level for men (N=34) was found to be 167.3 cm (SD 5.7) and for women (N=32) it was 157.9 cm (SD 5.6), but the adequate „pointing height” proved to be at practically the same level for both groups, i.e. at about the eye-level of the women. The „pointing height” was for the men 158.6 cm (SD 6.8) and for the women 157.1 (SD 6.9). A difference of only 1.5 cm, while the eye-level varied by 9.4 cm.

These results were controlled by experiments carried out with another 50 Ss, one half of which were women and the other half men. For the control experiments another method was employed where the Ss also had to determine their own eye-level. A special contrivance was attached to

## 3

Two types of means are of special interest when judging the results. The first of these is simply the distance with regard to the errors in pointing, irrespective of the direction or position in relation to the target. The second is the centre of gravity for the errors which, if the points of impact of the pointings were distributed quite at random around the target, would coincide

TABLE 1

	Male				Female			
	Right		Left		Right		Left	
	Dis- tance	S. D.	Dis- tance	S. D.	Dis- tance	S. D.	Dis- tance	S. D.
Luminous point . . . . .	15.5	8.2	15.0	7.5	22.3	14.3	24.8	19.1
Pointing with shut eyes. . . . .	26.0	8.3	24.8	10.7	30.9	12.3	33.4	14.6
Pointing from below . . . . .	17.1	7.5	17.9	10.6	27.6	11.3	25.1	10.9

the wall, and a small red ball, which served as the fixation point, could be moved up or down in a narrow groove by means of a handle that was placed at waist height directly in front of the S. The Ss were instructed to adjust first the fixation point at the height they found most „convenient” for observation. On this occasion not a single one of the 50 Ss experienced any difficulty in carrying out this test, and the natural “vision height” could be easily determined. The mean objective eye-level was found to be 168.4 cm (SD 4.9) for the men, and 158.0 cm (SD 5.1) for the women. The values for the „vision height” were 159.3 cm (SD 7.9) for the men, and 157.2 cm (SD 7.7) for the women. *Even in this case it was observed that the height chosen, showed the same tendencies to coincide with the objective eye-level for the women.*

Immediately after this experiment the S was requested to adjust the position of the ball on a level corresponding with his objective eye-level. There was no time restriction, and the S had ample opportunity to make this adjustment as carefully as possible. As a result of this test it was found that for the male Ss the estimated height was, on an average, 7 mm below the objective eye-level, while for the female Ss it was 7 mm above. The mean deviation from the objective eye-level, irrespective of the direction, was 2.2 cm for the men and 2.4 cm for the women — an error that seems to be surprisingly great with regard to the favorable conditions of the experiment. The 7 mm above and below the correct height thus amount to about 30 % of the deviation error.

(What is here only briefly referred to, is more fully treated and discussed at length in Sandström, op. cit. Ch. II).

with this. As will be shown below this does not occur in a single case.

Table 1 gives the means of the distances in mm for the different pointing experiments, with the exception of the variations of the luminous point experiments where the pointings were only carried out with the right hand. In the first two experiments under consideration, there were 34 male Ss and 32 female Ss, while in the pointing-from-below experiment there were 23 men and 25 women. The deviations of the female Ss are seen to be considerably greater throughout than those of the male Ss, and the greater errors made by the women as compared with those made by the men, were for the right hand, when expressed as percentages: 43.5, 18.9 and 61.7 %. While for the left hand these percentages were: 65.3, 34.9 and 39.9 %.

Table 2 shows the mean distances in mm for the above-

TABLE 2

	Male				Female			
	White light		Red light		White light		Red light	
	Dis- tance	S. D.	Dis- tance	S. D.	Dis- tance	S. D.	Dis- tance	S. D.
Dark-adaptation .	14.3	6.6	15.3	5.6	14.6	8.0	16.2	8.0
Without dark- adaptation . . .	14.9	10.3	15.0	7.5	18.2	12.7	19.5	15.4

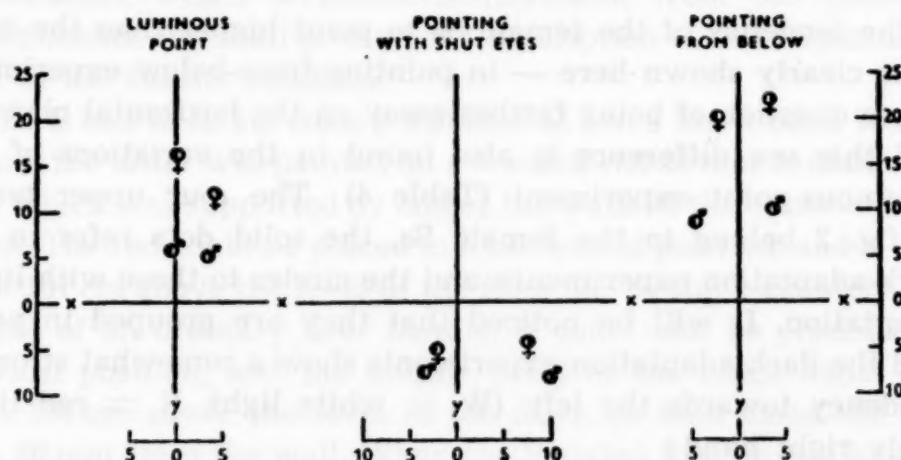


Fig. 1 shows the centres of gravity for the different pointing experiments. ♂ ♂ = male, left hand and right hand, ♀ ♀ = female, left hand and right hand. Distances in mm. (Scale 1 : 1)

mentioned variations of the luminous-point experiment. It will be observed that the dark-adaptation has almost entirely effaced the sex difference in connection with pointing to white and red light, but these still persist in the experiment without dark-adaptation. The percentage of greater error for female Ss was here 22.2 and 29.6 %.

If we take into account the centre of gravity relationships, we shall find that a sex difference of a specific type occurs without exception. Table 3 shows the points of gravity on the x and y axes for the experiments referred to in table 1, while fig. 1 elucidates the relationships with regard to position.

TABLE 3

		Right hand				Left hand			
		x-axis		y-axis		x-axis		y-axis	
		C. G.	S. D.	C. G.	S. D.	C. G.	S. D.	C. G.	S. D.
Male	Luminous point	+ 3.4	6.9	+ 5.1	11.2	- 0.5	9.8	+ 5.8	9.6
	Pointing with shut eyes. . .	+ 10.1	11.1	- 8.0	14.5	- 3.6	13.5	- 7.5	17.2
	Pointing from below . . . .	+ 3.8	9.1	+ 10.0	10.4	- 4.7	12.4	+ 8.8	13.1
Female	Luminous point	+ 4.1	11.4	+ 12.0	17.7	+ 0.3	14.9	+ 16.4	20.4
	Pointing with shut eyes. . .	+ 7.6	15.0	- 4.3	22.2	- 2.0	18.4	- 4.9	26.2
	Pointing from below . . . .	+ 3.5	15.3	+ 22.0	11.4	- 2.1	13.0	+ 20.3	11.5

The tendency of the female Ss to point higher than the male Ss is clearly shown here — in pointing-from-below experiment it is a question of being farther away on the horizontal plane — and this sex difference is also found in the variations of the luminous-point experiment (Table 4). The four upper points in fig. 2 belong to the female Ss, the solid dots refer to the dark-adaptation experiments, and the circles to those with light-adaptation. It will be noticed that they are grouped in pairs, and the dark-adaptation experiments show a somewhat stronger tendency towards the left. (W = white light, R = red light. Only right hand.)

From the results that have now been given, it is seen that the female Ss in general make greater mistakes in this kind of pointing than do the male Ss. Furthermore they point



TABLE 4

		White light				Red light			
		<i>x</i> -axis		<i>y</i> -axis		<i>x</i> -axis		<i>y</i> -axis	
		C. G.	S. D.	C. G.	S. D.	C. G.	S. D.	C. G.	S. D.
Male	Dark-adapt. . .	- 5.8	7.6	+ 0.9	11.5	- 6.9	6.2	- 0.8	12.9
	Without dark-adaptation . .	- 1.7	9.5	+ 3.1	13.7	- 0.1	8.9	+ 4.5	12.0
Female	Dark-adapt. . .	- 4.4	5.9	+ 7.7	11.2	- 5.1	6.4	+ 7.9	12.7
	Without dark-adaptation . .	- 2.7	6.8	+ 9.2	17.9	- 1.5	7.9	+10.1	20.2

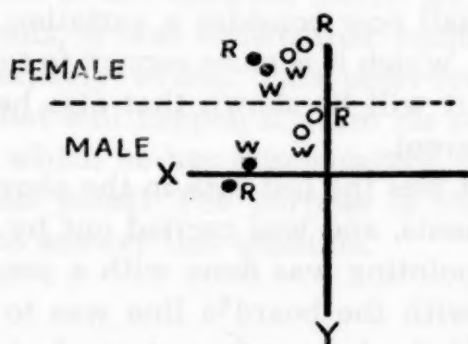


Fig. 2 (Scale 1.5 : 1)

higher than the men, and this tendency is apparent in all the experiments described. This also holds good for the experiment mentioned in note 4. This tendency is also evident in another experiment, which is somewhat different from our pointing experiments. We shall give a brief description of this experiment and of the results obtained.

The S had to fixate from a distance of 2 m a black band 70 mm broad; the latter was painted on a wooden rod 20 mm in diameter, and which was supported by one of the walls of the experimental room. The rod could be placed in a horizontal position, and it was then 107 cm above the level of the floor, i.e. at approximately the height of an ordinary door handle. It could also be placed in a vertical position, and the middle point of the black band was then 136 cm above the level of the floor. In both cases the rod was 60 mm from the wall. When the S stated that he had looked at the black band for a sufficient length of time, the E turned out the light, and in complete darkness the S stepped forward to grip the band by covering it exactly with his hand.

Also in this experiment the results showed that the female Ss made greater errors than the male Ss for both the horizontal and vertical positions of the rod. In the vertical position the centre of gravity for the male Ss was 2.6 mm above for the right hand, and 2.7 mm for the left hand. For the female Ss the respective figures were 15.1 and 14.7 mm <sup>5)</sup>.

## 4

All the tasks that have been described so far, may be termed localization experiments. It was a question of determining and indicating the position of a point in space, under certain given conditions. We shall now consider a variation of the luminous-point experiment, which it is more correct to term an orientation experiment, and it will be shown that also here a striking sex difference is apparent.

The experiment was the last task in the above-mentioned first series of experiments, and was carried out by all the 66 Ss.

This time the pointing was done with a pencil; and from the point of impact with the board a line was to be drawn to the luminous point. If the latter was not reached immediately, the S was to continue to try and find it without removing his pencil from the paper. Before starting, the task was looked upon as being very simple — no one expressed any doubts that complications might arise.

We mentioned in section 1, that it was typical of nearly all the Ss to feel quite assured that they could definitely locate the point of impact in relation to the target. Exact statements were made such as "2 cm up to the left", "1.5 cm directly below", "1 mm above to the right" etc. As a rule, when the light was

---

<sup>5)</sup> The characteristically similar values obtained for both the right and left hand are of interest. The same was observed in an experiment carried out by Stock (Z. Sinnesphysiol. 64 : 229—250, 1933), which it is interesting to compare with our experiments. She determined the degree of accuracy with which the S, when a point in space was tactually perceived with one hand on one side of the median plane, could indicate with the other hand the symmetrical point on the opposite side of the median plane of the body. She found that there were great individual differences in this ability, *but no difference was observed as regards the accuracy of the right or left hands*. The same was found to be true in our experiments for both the male and female Ss.

obscured by the S's hand, he or she believed that the target had been hit. In general, how is it possible to determine a position in this way? The S was convinced that he could point to the target, and when this proved not to be the case, what sense organ could possibly inform him of the error made? As far as we understand there is no such possibility, and there can be no question of knowledge that is based on any kind of sense data acquired.

The strength of the conviction of this ability to indicate the position, appears to depend on the vital need for the organism to be oriented and to maintain a valid connection with the actual frame of reference to the external world. At a very early stage of our experiments, it was shown that frequently the position indicated was entirely wrong. This observation led us to ask the question: What will happen if, from his point of impact, the true location of which he has misconceived, we let the S try to find the luminous point? The purpose of the experiment just described, was to answer this question.

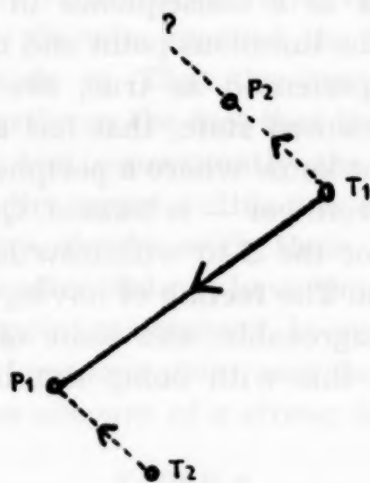


Fig. 3

Fig. 3 illustrates schematically a situation of this kind. Pointing has been carried out and the new experimental moment is to commence.  $P_1$  is the objective location of the luminous point, and  $T_1$  is the objective location of the point of impact, and the continuous line between these two points represents the path and the direction, necessary to take in order to come directly to the luminous point.  $P_2$  is the location that the luminous point is supposed to occupy in relation to the objective

point of impact, and  $T_2$  is the subjective location of the point of impact. The dotted line  $T_2-P_1$  is the line along which the S believes — more accurately “knows” — that he is moving his hand; but in reality he draws the line  $T_1-P_2$ . When  $P_2$  is reached and the luminous point still continues to shine, a critical situation must evidently arise.

Faced with the sudden discovery that the luminous point was not situated where it, with such confidence, had been assumed to be, the individual has naturally to reconsider the new situation. This process usually leads to the complete loss of any criteria, and the validity of the luminous point as a relational point no longer exists. This condition implies that the state of being oriented can no longer be maintained and is followed by a disoriented state. From a technical point of view it is interesting to observe that the oriented state was lost without any objective change having taken place in the stimulus field. We have carried out some movements with our right hands in a purposive way; these movements have not led to the expected result, and as a consequence of this, the relational conditions between the luminous point and the forefinger, which were previously experienced as true, are now resolved. The reason for this disoriented state, that has arisen, is apparently that we simply do not know where a peripheral part of our body — in this case the forefinger — is located. Quite a usual reaction to this situation is for the S to withdraw his hand and to start the experiment again. The feeling of having gone astray is often very strong and disagreeable, and some of our Ss have spontaneously compared this with being lost in a wood or in the wilds <sup>6)</sup>.

## 5

When considering the results, certain difficulties naturally arise as regards finding proper norms for what is meant by the oriented state actually experienced. This state cannot be judged according to the linear system that is a result of searching for the luminous point. At times this may be quite simple, and

---

<sup>6)</sup> Descriptions of the way in which the Ss react and figures showing the linear patterns executed when searching for the luminous point will be found in Sandström, *op. cit.* p. 146 ff.



the time taken very short, and yet a clearly expressed state of being disoriented may have prevailed. On the other hand the S may have worked for a comparatively long time, and his linear pattern may have been very complicated, without practically any sense of lost oriented state having been alluded to. It is evident that the individual's ability to withstand the stimuli that are liable to disintegrate his oriented state, are very differently developed.

The only possible course appears to be, to base our judgements on the reactions described by the Ss during and after the trial, even if in certain cases these may be difficult to interpret. That spontaneity and suggestibility are possibly more highly developed in women has been taken into account, and borderline cases have been taken as belonging to those who maintained an oriented state.

It would appear that a natural rough classification of the results would be to place the Ss in two categories: one for those who remained oriented, and the other for those who became disoriented. These two categories have then been subdivided into two groups: the Ss who reached the luminous point, and those who failed to do so. That the luminous point was not reached depended partly on the fact that the S's hand concealed the luminous point, and consequently the S assumed that he had already reached the target — this was usually the case with Ss who worked very slowly with their pencils. Those who worked very quickly often did not have time to stop their hands before the luminous point reappeared. In several cases, however, *failure to reach the luminous point was due to the experiment being discontinued on account of a strong feeling of discomfort.*

TABLE 5

	Male (N = 34)		Female (N = 32)	
	Oriented	Disorient.	Oriented	Disorient.
Reached luminous point . . .	14	4	4	13
Did not reach luminous point .	9	7	7	8
Total . . . . .	23	11	11	21

Table 5 shows clearly to what a much higher degree women are susceptible to stimuli that have a disintegrating effect on the oriented state. 21 out of our 34 female Ss stated that they had

lost their sense of being oriented — about  $\frac{2}{3}$  of the material — while only 11 out of a total of 34 male Ss reported that they had been affected in the same way, i.e. practically  $\frac{1}{3}$  of them. (Significant at 1 % level.) This difference is particularly marked for those who accomplished the task, while the figures are very similar for the Ss who were unable to find the luminous point. The achievements in this row are more difficult to judge with regard to the factors just referred to above, which in different ways were the cause of the failure to reach the luminous point. It is noteworthy, however, that so many of the male Ss who became disoriented were unable to complete the task.

## 6

In this paper we have described the results obtained in a more comprehensive investigation; and moreover we have considered the findings exclusively from the point of view of sex differences.

We should like, however, to say a little more about the difference between localization and orientation alluded to above. There are good reasons for restricting ourselves to operational definitions of the two concepts, even though it seems probable that our definitions are, to a certain extent, also applicable beyond the range of the experiments described.

Localization refers to the process by which we determine and indicate the position of an object in space. This implies the position of an object in relation to the subject.

Orientation, in the strict sense of the term, refers to a process whereby we pass from a disoriented state to an oriented one. This process differs in the most essential manner from localization, inasmuch as it is here a question of determining the position of the subject to an object. This is experienced very strongly in our orientation experiment described above. When the oriented state is lost, a moment of urgent need intervenes in the situation, and an important characteristic of a need is its personal centeredness, and that under the pressure of need the direction of the modes of behaviour is from within outwards. These modes are active and not reactive. The loss of the oriented state compels us to endeavour to re-establish this state as quickly as possible; and until this is achieved we experience a strong sense of discomfort and unrest. In our use of the term, orientation signifies an active process; while localization is a reactive

process that has nothing to do with the satisfaction of needs. Localization presupposes a state that is experienced and conceived of as being oriented.

Orientation ability is, in accordance with our line of argument, the ability to pass from a non-oriented to an oriented state. As far as we are aware, no investigation has been undertaken of this ability in the sense in which it is used here. What has been investigated is *the ability to remain oriented*. When we speak, for example, of the highly developed orientation ability of primitive races, we refer to their ability to resist stimuli that tend to disintegrate the oriented state.

In a wider sense, orientation ability may be defined as the ability to receive external impressions, and to interpret them in such a way that at every moment, they are integrated in a valid frame of reference. By these means, the oriented state is maintained continuously.

From our point of view, the decisive criterion, that makes searching for the luminous point an orientation experiment, as distinguished from the other experiments already described, is that the character of the experiment is such, that it can give rise to the loss of the oriented state. It also appears to be of importance here that no sense organ that is relevant to the situation, is artificially prevented from functioning — above all, the Ss' eyes were not blindfolded — and the darkness that surrounds the luminous point is not regarded as an obstacle in performing the task.

In accordance with these definitions, it appears to be evident from our results that women have a less developed ability to localize an object than men have; and the same holds good with regard to their ability to withstand stimuli that are liable to disintegrate the oriented state. The orientation ability in the strict sense of the term as used above, from the point of view of sex differences, is more difficult to determine. Perhaps the results given in table 5 may be taken to indicate that the ability of the women is greater here since so many of those who had lost the oriented state succeeded in re-establishing this state by actually reaching the luminous point. Among the male Ss who became disoriented, there were fewer who managed to reach the luminous point than those who failed to do so. The material hardly permits of an interpretation of this kind on account of

the influence of many factors that were not sufficiently under control; but we are anxious to emphasize that the contrary interpretation is also precluded.

## 7

In general it may be said that the aspect of sex differences that is dealt with here, has been given very little attention. Out of all the investigations where Ss were required, when blindfolded, to go to, or in some other way to approach a given target — this type of tests deserves to be called the classical orientation experiment — only Szymanski has mentioned sex differences <sup>7)</sup>. He found that the girls who formed part of his material had greater deviations. In as far as the results are relevant to orientation in our sense, McNemar <sup>8)</sup> points out that the items involving 'orientation' in the Terman-Merrill test scale show superiority for boys in space ability.

As regards localization experiments, it is apparently possible to state that the women showed a lower degree of sureness of aim than men did, and thus a condition is indicated that was observed previously, and which was also demonstrated, *inter alia*, by Whipple and later by Gesell for different age groups <sup>9)</sup>.

The research work, however, that appears to us of particular importance, is that which has been carried out by Witkin during the last few years. In 1948 he published together with Asch the fundamental exposition, and what they particularly investigated was the perception of the true upright and horizontal <sup>10)</sup>. In later papers <sup>11)</sup>, Witkin has, *inter alia*, shown that significant sex differences are found in perceptions of that kind.

<sup>7)</sup> Some of them are Guldberg, F. O., *Z. Biol.* 35 : 419—458 (1897), Lund, F. H., *Amer. J. Psychol.* 42 : 51—62 (1930), Schaeffer, A., *J. Morph.* 45 : 293—398 (1928) and Szymanski, J. S., *Pflüg. Arch.* 151 : 158—170 (1913).

<sup>8)</sup> McNemar, Q. *The Revision of the Stanford-Binet Scale*. Boston: Houghton Mifflin 1942, p. 52.

<sup>9)</sup> Whipple, G. M. *Manual of Physical and Mental Tests*. Baltimore: Warwick & York 1921. Gesell, A. *et al.* *The First Five Years of Life*. N.Y.: Harper 1940.

<sup>10)</sup> Asch, S. E. & Witkin, H. A. *J. exp. Psychol.* 38 : 325—327, 455—477. Witkin, H. A. & Asch, S. E. *J. exp. Psychol.* 38 : 603—614, 762—782 (1948).

<sup>11)</sup> Especially in Wapner, S. & Witkin, H. A. *Amer. J. Psychol.* 63 : 385—408 (1950), Witkin, H. A. *Psychol. Monogr.* 63, No 7 (1949), *Trans. N. Y. Acad. Sci.* 12 : 22—26 (1949) and Witkin, H. A. & Wapner, S. *Amer. J. Psychol.* 63 : 31—50 (1950).



## DAVID KATZ

David Katz, one of the greatest psychologists of our time, and professor in The University of Stockholm suddenly passed away on the second of February of this year. To me, his death means the loss of one of my oldest and most loyal friends. We studied in Göttingen under the guidance of G. E. Müller and took our degrees there at the same time in 1906. If now, after half a century, I recall those happy and, for both of us, decisive years of our lives, I can see how much we gave to each other and how those golden days were to determine the development and direction of our scientific work. In our thoughts, our feelings and intentions we have remained faithful to our science and each of us was happy to see how great a part the other played in developing and changing psychology.

David Katz was born in Kassel on the first of October 1884. First he studied mathematics and physics; it was only after some time that he turned to experimental psychology. In 1907 he became assistant at the Psychological Institute of the University of Göttingen, and in 1911 Privatdozent in psychology. In 1919 he was elected to the chair of psychology and pedagogics in Rostock. The Nazis deprived him, together with other eminent psychologists such as William Stern, Gelb, Selz, of his office in 1933. He was asked by Professor Pear to lecture at Manchester University and was there from 1933 to 1935. He then worked at London University College until 1937, when he accepted the chair of psychology and pedagogics at the University of Stockholm.

David Katz was a very active scientist, who worked in diverse fields with great success. His works appeared in a good many languages and had a profound influence everywhere. He was especially interested in *conceptual psychology*. His most important works, "Die Erscheinungsweisen der Farben" (1911) and "Der Aufbau der Tastwelt" (1925), as well as his investigations on the "Vibrationssinn" and on hunger and appetite, and even his "Gestaltpsychology" move mainly within the field of the psychology of concept, for which he was especially well-equipped through his great inventive and experimental genius and his sense of precision. He was very successful with his

excellent work "Animals and Man" (1937). Many of the experiments on animal psychology reported in this book we carried out together in Göttingen, and some of them we published in "Nachrichten der Gesellschaft der Wissenschaften", Göttingen and some in "Zeitschrift für Psychologie".

As research worker and organizer he played an eminent role in the international psychological movement. He was known everywhere, and everywhere his scientific advice was in demand. He was frequently honoured for his very fine work by expressions of appreciation from his contemporaries. He was President of the Swedish Society for Psychology, honorary member of the English, French and German psychological societies, secretary of the International Union of Scientific Psychology, co-editor of "Acta Psychologica" and of "Zeitschrift für Psychologie" and corresponding Member of the Bavarian Academy of Sciences. He was also the President of the XIIIth International Psychological Congress in Stockholm.

David Katz was a very serious and conscientious scientist, but he also had a very strong sense for the humorous situations of life. Everywhere he was appreciated and loved. His pupils will always value the seeds he disseminated and marvel at his scientific ethics. We, his friends and colleagues, will never forget him. Through his qualities both human and as a scientist, his personality always made a unique impression on us. Our greatest sympathy goes out to his wife, Dr. Rosa Katz, who through the death of our friend, will have the greatest loss to bear; she has been not only a devoted wife but his loyal companion in his scientific work. Our friend Katz left two gifted sons: the elder is a scholar and a teacher of English and German in Stockholm, the other a medical practitioner.

Gradually the older generation of psychology leaves us. The younger generation of psychologists must prove themselves worthy of their inheritance.

G. RÉVÉSZ.

*Laboratory of Personality Assessment & Group Behavior,  
University of Illinois*

A QUANTITATIVE ANALYSIS OF THE CHANGES IN THE  
CULTURE PATTERN OF GREAT BRITAIN 1837—1937,  
BY P-TECHNIQUE

BY

RAYMOND B. CATTELL

1. THE DEVELOPMENT OF FACTORIAL TECHNIQUES IN SOCIAL SCIENCE

A proper appreciation of the findings of this study requires that the reader have some familiarity with certain new methods and concepts used in a whole series of researches on culture pattern dynamics (9) (11) (13) (14) (26) (26a) and in experiments with small groups (9) (10). A brief outline of these methods and concepts can, however, be presented in a few paragraphs to the reader versed in elementary statistics.

Our first tenet is that the proof of functional connections in sociological, anthropological and historical phenomena must rest on demonstration of correlation. Many anthropologists and most historians are still compelled to proceed to hypotheses and explanations without this kind of statistical demonstration, though it may be conceded that they use intuitive "mental calculations" which assume methodologically the same form. Anthropology in particular has remained immature as a scientific discipline through failure either to work out reliability coefficients between observers regarding the raw observations, or to demonstrate positive and significant correlations among the alleged elements of a culture pattern, despite the early examples set in methodology by the work of Hobhouse, Wheeler and Ginsberg (24) Unwin (40) and a few others.

Perhaps this neglect of rigorous procedures stems partly from a feeling that the demonstration of relation between only two variables is no more capable of picking up the total pattern in which the anthropologist is interested than chop sticks are capable of picking up soup. This objection is important, but is

met by our second tenet: that analysis in social science needs to operate simultaneously with *a wide and strategically chosen range of variables* rather than with the paired "dependent and independent variable" of classical physical science experiments. This can be achieved at present only by applying to the correlations the method of factor analysis (15) (38).

The concepts which march with this method are (1) that a group specifically defined as *an aggregate of persons in which all are necessary to some satisfactions of each*, has some properties of an organic unity. For example, reliability coefficients obtained for the performances of experimental groups (17) as well as for nations over two fifty year periods (13) are positive, significant and of about the same order as for individual personality traits; (2) that the description and measurement of a group requires attention to three distinct panels (9) of data: (a) population data i.e. *means* of traits for all *individuals* in the population (b) structural data, i.e. institutions, as *relationships* of individuals, and (c) syntality data, traits of the total organic group inferred from the performance of the group as a group (3) that syntality may be measured in terms of numerical values for a relatively limited set of dimensions, abstracted from actual group performance by factor analysis of a wide variety of behavioral variables.

A number of theorems have been developed (9) and recently checked by experiments on a hundred small groups (17) concerning some of the above, notable as to the nature of the functional unities constituting syntality dimensions, the use of individual syntality profiles in the specification equation (15) the relation of syntality to synergy (the total interest going into a group) and so on (17). But as background for the present article we need enlarge only upon the findings on syntality dimensions and culture patterns for national groups.

Seventy-two "culture pattern" variables — economic indices, vital statistics sociological, psychological and historical measures (pertaining to 1836—1936) — have been factor analyzed for a population of sixty-nine countries, yielding twelve significant, independent dimensions for measuring a culture pattern, as described elsewhere (11). A recent, more intensive study of forty countries (16) has confirmed eight of these factors. When all contemporary national cultures are expressed as profile patterns



on these dimensions family resemblances are observed among them, the degree of which can be accurately assessed by the *coefficient of pattern similarity* (15). The application of this criterion has led to the discovery of seven or eight major "culture pattern families" agreeing (13) well with anthropologists intuitions and with the "civilizations" designated on historical grounds by Toynbee (39).

The particular factor analytic design used in discovering these unitary "traits" or dimensions common to all cultures is that known as R-technique, (15) which correlates variables for a *series of countries* i.e. operates on individual differences of countries. It was the success of R-technique in structuring the complex field of human abilities and personality traits (12) which actually suggested its application to the equally complex and chaotic realm of social and cultural variables. But personality study has recently gained additional powers of definition from a new development of factor analysis known as P-technique <sup>1)</sup> (8) (14) (43), and it seemed that it would be productive of new insights to apply this also to group data i.e. to nations as organic wholes which is indeed the purpose of the present pioneer study.

P-technique, instead of operating on individual differences among cultures, as R-technique does, takes a single culture and observes its changes over a period of time, under the impact of external stimuli and inner forces of development or regression. A series of variables is measured every year (or at some other interval) for the given country over a hundred years, or some period sufficient to provide a length of series capable of giving correlations with a small enough standard error for the purpose concerned. Thus if armament expenditures correlate negatively with degree of liberalism of parliament, or severity of discipline in schools correlates with expansiveness in politically controlled area (a led-and-lag correlation being used in the latter case) we argue that real functional connections, of a psychological or social nature, exist between the given variables. Such longitudinal correlations have already been used — frequently in economics, rarely in history — but their integration in the factor analytic method, and the using of all technical skills accumulated

---

<sup>1)</sup> P-referring to *process* factorization.

around P-technique, is apparently new. For P-technique, like any factor analytic method, is a wholistic method, which proceeds to uncover behind the variables the underlying massive influences, of which the variables are superficial operational indicators, and in this case we should expect the influences to be historically major developments or directions of change in the culture pattern.

## 2. THE RESEARCH DESIGN AND THE CHOICE OF VARIABLES

The aim of factor analysis is, first, to replace an almost infinite number of possible variables by a more compact list of basic dimensions in terms of which all kinds of comparisons, investigation of functional relations, etc. can be more economically transacted and, secondly, to produce meaningful hypotheses from inspection of the variables forming each factor cluster, as to the underlying influences at work. From the standpoint of pure science the second is more important, and our paramount objective in this study is consequently to obtain hypotheses as to the functional connections in historical forces.

However, factor analysis is a procedure of abstracting from particulars and it therefore seeks those generalized processes of social dynamics, beyond historical accidents, which operate in all culture patterns at all times, though in different degrees. This search is most effectively pursued when R-technique and P-technique are used in conjunction in a strategic research plan including several statistical independent but coordinated studies. For if a functional unity has more than transient reality it should appear both when examined in terms of describing individual differences and in terms of a unity of development i.e. a pattern in *intra*-individual differences. For example, Spearman's "g" shows itself in that an individual higher than another in spatial ability also tends to be higher in verbal analogies and classification performances (R-technique). But it can be demonstrated also in the observations that as a child grows older he improves simultaneously *all three* (P-technique). In personality research it has similarly been shown that the *common* factor pattern of the principal R-technique factors appears again, with slight individual modifications, in the *unique* factor patterns obtained from single individuals.

In the systematic approach of which this paper is part it is

important likewise, in the realm of national culture patterns, to see whether (1) the patterns discernible by R-technique can also be found by P-technique and (2) the pattern from P-technique on one country is similar to that from another. (The *mean* P-technique pattern one might expect to equal the R-technique pattern). To make possible a check on these hypotheses it is obviously necessary that the different researches should have a *majority of variables identical* from study to study. Consequently the choice of variables requires careful planning.

The original list of eighty variables in the R-technique studies (11) (16) was chosen on the basis of testing hypotheses there (11) indicated e.g. Unwin's theory of the functional connection of sex restriction and religion (40); on the basis of catholicity among the interests of the social sciences; but ultimately by the ruthless limits of availability of quantitative data for 69 countries. The present choice followed the same guides, but, in addition, met the need (1) that the variables should be repeatedly measurable each year, and (2) the above mentioned condition that the data could be found also in the R-technique lists for comparison, and (3) that they should be available for other countries on which later longitudinal studies might be made on the same lines.

Because of the almost unequalled completeness of records, as well as on account of its great interest to our own culture family, Great Britain was chosen as the country of study, but a comparative study with substantially the same variables was simultaneously undertaken for the U.S.A. as reported elsewhere (14). Since we might anticipate from six to a dozen factors it seemed desirable to have a "population" of not less than 100 years. To avoid the effect of the second world war and the incompleteness of more recent records, we took the period 1837—1937.

The followed list of 48 variables (Table 1) numbered as in the Factor Matrix (Table 2) has the adjectives "high", "low", "few" etc. attached to each, to indicate in which direction the variable was scored positive (high) for the purposes of subsequent correlation. To the left of each variable a letter R is placed if it was used also in the international R-technique studies (11) (16) and a P. if used in the U.S. P-technique study (14). The numbers attached to the letters are the matrix numbers

of the variables in those studies and permit one readily to discover how the variable behaved in these other structurations.

TABLE I

*List of Culture Pattern Variables Used in P-technique Analyses of Britain 1837—1937*

No. in Related Studies	No.	Title
R47	P1	1. High national debt (5)
		2. High number of ministries (6) (32)
		3. High average age of ministers (6) (32)
R10	P9	4. High gross population (32)
R56	P41	5. High density of population (32)
R9	P11	6. High gross birthrate (28)
R18	P15	7. High gross deathrate (2) (33)
		8. High proportion of civil to religiously affiliated marriages (3)
R	P22	9. High mean annual temperature at Greenwich (3)
		10. High number of criminal convictions (total population?) (33) (41)
		11. Left wing in political power (negatively scored when (5) conservatives in office)
R14		12. High trade union membership (per 100,000 population) (20)
R17		13. High number political parties in Commons (21)
R3		14. High severity of war (number of war casualties, gross) (21)
		15. High illegitimate birthrate (3) (32)
	P29	16. Total foreign trade per capita (35) (41)
	P14	17. High number of patents sealed (31) (33) (41)
R13	P30	18. Low favorableness of balance of trade (Excess of imports over exports) (35) (41)
R4	P3	19. High expenditure for defense (In pounds, gross) (6) (41)
R57	P12	20. High number of divorces and annulments (33) (41)
		21. Large number of acts passed to improve status of women (23)
		22. High favorable balance of budget. (Revenue minus expenditure) (5) (33)
		23. High marriage rate (Gross) (3)
	P40	24. High proportion of population registered parliamentary electors (22) (29)
R12		25. High percent professional occupations (32)
	P20	26. High percent domestic occupations (Census)
	P19	27. High percent commercial occupations (Census)
	P17	28. High percent fishing and agricultural occupations (Census)



TABLE 1 (Continued)

*List of Culture Pattern Variables Used in P-technique Analyses of  
Britain 1837—1937*

No. in Related Studies	No.	Title
R70	P18	29. High percent industrial occupations (Census)
		30. High proportion of fame is political (Percentage of cases in Who's Who) (42)
R48		31. High proportion of fame is scientific (Percentage of cases in Who's Who) (42)
R24, 58	P2	32. High emigration
	P31	33. High length railroads (22) (33)
		34. High rating for prosperity (27)
	P13	35. High number of families per house (41)
		36. High ratio of women to men (41)
R67		37. High number railroad passengers (33)
	P33a	38. High illiteracy — women (1) (3) Ratio female to male illiterates — R42
	P33b	39. High illiteracy — men (1) (3) Ratio female to male school attendance — P34
R22		40. High government expenditure for education (32)
R41	P21	41. High per capita alcohol consumption (33)
R6		42. High number deaths from alcoholism (3)
		43. High strength of police force (Policemen per 1000 pop.) (6)
R59	P38, 39	44. High number of deaths from cancer (25)
R32	P42	45. High number suicides (3)
R23		46. High number homicides (3)
R15		47. High number deaths from syphilis (3)
		48. High number deaths from typhus (3)

In the following cases the definitions of variables in the related studies were slightly different from those here, though sufficiently related for mention: 14 in R3 was "frequency of involvement in war"; 18 in R13 was reversed, as „ratio of value of exports to imports"; 20, in R57 was ratio of divorces to marriages; 25 in R12 was tertiary occupations, not merely professions; 34 appeared in R24 and R58 respectively as real income per head and real standard of living; 33 and 39 also appeared as R42 and R34 "ratio of female to male illiterates", as well as in the sum „Total Illiteracy" as R67.

The sources of data from which these variables were obtained over the hundred years are shown by the bibliographical references after each, and in several instances they were checked from more than one source or obtained from the collation of

material from sources too numerous to list. Though the majority were available at yearly intervals, some, notably the census figures, were available only for ten year points or ten year intervals. In these cases the points were plotted on a graph, a curve smoothed to the points, and the annual figures obtained thus by intra-polation.

### 3. THE COURSE OF THE FACTOR ANALYSIS

Since all of the above variables are expressed in some kind of numerical continuum the 1128 correlations of the correlation matrix were worked out by the product-moment formula.

The factorization was carried out by the *grouping* method (15), the extraction of twelve factors being required as judged by residuals. However, three of these factors were found in the rotation to be superfluous, and are reproduced here (along with the excessive communalities associated with them) only for the sake of a complete computational record. As usual the rotation for simple structure was carried out "blindly" i.e. by trial and error, without knowledge of the meaning of the variables and therefore *without prejudice from pre-conceived theories*. It will be seen from Table 2 that by ordinary standards the simple structure is a good one, showing, with two exceptions about two thirds of the variables (or more) in the  $\pm .10$  loading hyperplane. The resolution also looks rather unusually clean in the graphs.

However, there is an anomaly in factors 1 and 2, factor 1 having only 6 variables in the hyperplane. It is noteworthy that a similar condition was found in the American P-technique study (14), factor 1 having only 8 variables in the hyperplane. It would seem that in historical data there are always one or two trends so massive that very few variables indeed are left uninfluenced by them. Certainly every conceivable attempt was made to rotate factor 1 on other factors in such a way as to increase its hyperplane, but after an unusually thorough and prolonged rotation process, of 28 over-all rotations, it was clear that no further improvement was possible. Our confidence in the correctness of the present rotation of Factor 1 therefore rests on the definiteness and constancy of the present hyperplane, brief though it is, and on the essential orthogonality of this factor to all the other factors.

TABLE 2  
The Rotated Factor (Reference Vector) Matrix

No. of Variables (as in Table 1)	FACTORS										h <sup>2</sup>
	1	2	3	4	5	6	7	8	9	10	
1	26	08	36	-01	08	00	02	-13	18	26	.848
2	03	-43	02	-11	12	08	-01	-04	08	01	.715
3	17	-23	20	-03	05	44	04	-02	02	08	.326
4	44	-16	06	00	01	-04	05	-04	-09	-03	.984
5	44	-16	06	01	-00	-04	04	-05	-09	-03	.981
6	-49	-10	-20	04	04	10	00	02	-11	00	.917
7	-45	05	-08	-02	01	12	03	04	-00	01	.925
8	48	-10	10	03	04	-63	05	-06	-09	-01	.972
9	37	-07	09	-16	05	-07	-02	-02	-02	-05	.257
10	-48	19	-10	-05	01	-12	-08	09	25	-01	.919
11	-23	-07	12	05	12	-21	07	-01	-08	-07	.158
12	27	09	13	00	01	-25	64	-06	01	05	1.579
13	47	-01	02	05	-05	-06	-10	-09	-00	-01	.784
14	10	45	-09	-05	-24	-00	-05	04	-01	09	.382
15	-44	12	-08	01	09	-02	06	-00	02	04	.962
16	50	02	06	06	07	-14	09	-03	02	04	.737
17	44	-14	08	-07	03	01	-07	01	-06	-03	.937
18	72	33	16	-05	-01	08	-01	-01	02	10	.889
19	65	53	-05	04	11	-02	-03	05	-08	-01	.818
20	12	01	31	-07	-03	-10	03	-63	05	13	1.563
21	48	-20	23	02	-01	05	-02	-10	-04	08	1.006
22	-61	-55	07	-04	-04	-04	04	-02	09	-01	.781
23	13	18	-01	-21	27	02	18	-14	-01	13	.312
24	17	-01	94	-28	-05	-03	-02	10	03	16	1.526
25	48	-08	-00	-17	-10	01	05	07	-14	01	.964
26	-51	-34	45	09	09	08	08	06	-12	00	.909
27	22	-12	50	04	05	-07	-02	-11	-09	12	.663
28	-41	18	02	-01	04	00	01	-02	28	04	.960
29	45	-10	-60	23	-01	04	-21	04	-25	-15	1.251
30	44	-22	-02	01	13	11	-39	-07	-28	-06	1.106
31	-08	-16	-04	40	-14	03	08	-03	-24	-04	.581
32	-12	-11	-11	38	07	-55	-00	03	-05	-30	.671
33	36	-07	02	-08	-07	-08	02	11	-13	-09	.876
34	-02	03	-12	-04	16	-06	00	-01	01	-06	.092
35	-38	-04	-04	-10	42	-02	32	-60	-13	12	1.360
36	56	-04	22	-04	07	07	10	-09	07	14	.871
37	54	-05	09	-01	03	-05	05	00	-09	-00	.990
38	-40	-01	06	16	14	02	01	-14	08	05	.948
39	-43	18	02	04	03	-05	-04	-02	23	03	.959
40	48	11	27	-08	07	-14	05	-04	17	11	.985
41	01	08	11	-06	-28	22	-00	-18	-22	25	.546
42	-05	-34	-21	-02	-13	46	-19	07	-13	-04	.885
43	49	17	00	02	-13	-20	02	08	-05	-06	.925
44	67	17	10	04	-13	-03	-04	-06	-02	09	1.005
45	50	06	09	-03	-20	-07	-08	-02	-01	09	.956
46	-52	-17	-07	-02	36	-01	12	07	-01	-09	.938
47	-22	-40	-02	27	29	01	-11	09	-54	-19	.834
48	-21	-25	-01	66	-01	-01	-21	-03	13	-10	1.110
±.10 Hyperplane	6	21	30	38	32	36	39	40	13	38	

Although twelve factors were initially extracted the rotation reduced two of them to undoubted residuals (originally listed

as 11 and 12) while three more (Factors 8, 9 and 10 in Table 2) eventually showed such trivial loadings that we shall not attempt to interpret them in the following discussions. (Factor 8 has significant loadings only on variables which happen to have been estimated with too high a communality). Thus we have essentially seven factors, comparing with six in the American Study, the extra factor being perhaps due to the slightly longer series, slightly greater completeness of records and the somewhat larger and different population of variables.

The correlations among the factors (shown only for the seven used: the rest are normal) are set out in Table 3. It will be observed that in general they are quite small and insignificant, though in four instances they exceed 0.30, which is unusual in

TABLE 3  
R<sup>2</sup> Matrix: Correlations Among Factors (Reference Vectors)

	1	2	3	4	5	6
2	29					
3	-14	-07				
4	23	-23	-37			
5	-01	-02	20	-27		
6	27	-31	-09	00	-34	
7	-15	-42	-18	-12	-21	30

simple structure rotations. Since an undue number of definite, substantial inter-factor correlations was also found in the American P-Technique study we may suspect that factors in historical material are more likely to be intercorrelated than in other domains of use of this method. The particular correlations are discussed below.

The remaining statistics necessary for further examination or re-interpretation of these results, namely, the correlation matrix, the unrotated factor matrix and the transformation matrix are available on microfilm from the American Documentation Institute, Science Service Bldg., 1719 N. Street, N.W., Washington 6, D. C. under the reference number 3432.

#### 4. INTERPRETATION OF THE FACTORS

We shall arrange the evidence from which the nature of each factor may be tentatively inferred in the usual way, namely by picking out that small fraction of the total variables — about an



eighth — which is significantly and conspicuously loaded therein. In each case the verbal description' of the variable will have "high" or "low" attached to it to conform with the sense of the algebraic sign of the given loading.

FACTOR 1  
Cultural Pressure

Variable No.	Loading	Description
18	.72	Unfavorable Balance of Trade (More imports than exports)
44	.67	High death rate from Cancer
19	.65	High expenditure for defense (armaments)
22	— .61	Defective Balancing of the Budget (More expenditure than revenue)
36	.56	High ratio of women to men in the country
37	.54	High number of railroad passengers per annum
46	— .52	Low frequency of homicides
26	— .51	Fewer persons in domestic service occupations
45	.50	High frequency of suicide
16	.50	High total foreign trade per capita
21	.48	Many acts passed to improve status of women
17	.44	Patents for many inventions sealed

Because of the large variance in this factor a rather large number of variables are taken, for they are still all above a loading of .43. This factor is evidently one of crisis, with economic and psychological stresses, in part connected with outside threat of war — but it is not loaded in the variables of war itself. For example, No. 14., Casualty Rate, is only .10, and war as such is indeed clearly accounted for by factor 2 below.

Variables 44, 45 and 16 here, as well as something resembling 21 (U.S.P. 34), occur in the same pattern in factor 1 of the U.S. study (14), but 18 and 44 also occur in factor 3 thereof. As pointed out in that study factors 1 and 3 are substantially correlated, so our present factor could be a second order factor relative to both of them, but its very restricted hyperplane i.e. its excess of loaded variables, places it in the same relation to the total set of variables as factor 1 held in the U.S. study. With all these considerations we are compelled, therefore, to favor matching factor 1 here with factor 1 there. Now factor 1 there was labelled "cultural pressure" because of its resemblance to the R-technique factor (11) (16) so interpreted, but it is unmistakable that the remaining variables here give a distinctly

different twist to the meaning from that of the previous cultural pressure factors.

The cultural pressure factor has typically shown complication of life, signs of frustration, inward and outward aggression (suicide and war), and cultural productivity — with urbanization as a concomitant but not essential feature. Here however, the cultural productivity cannot be too clearly shown, because not included on the original variables. The frequency of inventions, however, which may act as our one true representative of the cultural productivity variables, has its one substantial loading in this factor and nowhere else. Perhaps we can also include reduction of homicides and improvement in the status of women as true manifestations of cultural pressure. The complication and aggression of the culture pressure concept exist in the present pattern but so also does economic stress — of the country as a whole. (Note full employment, low homicides but unbalanced budget). One is reminded in contemplating this complication, of the pattern of "anomie" asserted by Durkheim to be the matrix of high suicide (22), as well as of Znaniecki's concept of disorganization in the life of the immigrant (36). The "long circuiting" and complication of satisfaction i.e. the psychological concomitant of "anomie" which we have regarded as central to the "cultural pressure" factor can evidently have its origins in somewhat different sources. In general (R-technique study (11) (16)) this disorganization appears as a complication of culture as a whole (presumably by borrowing and invention); in the U.S. study it has distinct associations with urbanization and increase of population density, and in Britain it seems to be associated more with economic stresses of unbalanced trade and maintaining political position in the world. These issues will be given that fuller discussion which they require in a separate article (14), but we may venture that differences in the life cycle of the countries concerned could account for the different associations, i.e. causes and manifestations, of the essential stress-complication conditions.

Two of the three variables (18, 19, 26) common to this research and the U.S. study occur in factor 2 here and factor 2 there, suggesting a matching of these factors. (Factor 2 in the U.S. study (14) happens, however, to be inverted in algebraic sign). What we have called "Restrictive Hard Times" there takes on

## FACTOR 2

## War Stress — vs — Ease of Living

Variable Loading No.	Description
22	—55 Defective Balancing of the Budget
19	53 High expenditure on defense (armaments)
14	45 High casualties in war (severity of war)
2	—43 Fewer Ministries maintained in government
47	—40 Fewer deaths from syphilis
26	—34 Reduction of people in domestic service occupations
42	34 Fewer deaths from alcoholism
18	33 Unfavorable balance of trade

more obviously the character of "War Stress" here. (The fewer deaths from syphilis is anomalous unless we consider that army medical services actually reduce syphilis deaths despite increased possibilities of infection). Reverting to the U.S. pattern we notice now that "many men in the army per capita," loaded 0.55, should perhaps have indicated War Stress as a better interpretation of that factor also.

Now that this factor is located by a definite loading pattern future research can turn to the interesting task of concentrating on the more purely psychological variables which are thought to change with war and to be generated by the stresses of war.

## FACTOR 3

## Emancipation — vs — Rigor

Variable Loading No.	Description
24	94 High proportion of population registered parliamentary electors
29	—60 Reduction of number of persons in industrial occupations
27	.50 Increased number in commercial occupations
26	.45 Increased number in domestic occupations
1	36 High national debt
20	31 Increased number of divorces and annulments

Variables 24, 26, 1 and 20 occur in the top variables in Factor 3 in the U.S. study, and nowhere else therein with the proper sign pattern. Low birthrate, which occurs in the U.S. pattern, also has a substantial loading, —.20, here, but not enough to be in the top six. The matching of the two factors is rendered less perfect by certain variables here — 27, 29 and others somewhat lower — occurring in Factor 1, not 3, in the U.S. data. But this

is resolved by the fact that 1 and 3 are highly cooperative factors there whereas they are not so here.

What is common to both 3 factors is an increasing participation of the common people in government, together with amelioration of living conditions, increase in the national debt and decline in family stability. In the American setting this is associated with urbanization, increase in mental disease and increase in luxury consumption, whereas there is no proof or disproof of this aspect of the pattern here, through lack of inclusion of the necessary variables. If we take "ratio of divorces to marriages", "urbanization", "increase in tertiary occupations", and "growth of political preoccupations" (perhaps also frequency of telephones and industrial unemployment arises) as hallmarks of this factor we can recognize it in four separate researches: the two P-technique factors above and the R-technique factors (11) (16) which have been called "Pace of Life and Emancipation — vs — Unsophisticated Stability".

For this report we shall call the factor in its present setting "Emancipation — vs — Rigor", since the very large loading (0.98) specifically in political emancipation is to be discounted somewhat on account of a bad communality estimate on that variable. Further research should ask whether some variables here left in Factor 1, notably urbanization, mean annual temperature, and others shared with Factor 3 in the U.S. study might not here also be rotated into Factor 3. Additional crucial variables should then be introduced to decide among the alternative hypotheses that the positive sense of this dimension is either political democracy (as opposed to stable authoritarianism), or urbanization or relaxation of exacting, perhaps narrow, behavior standards.

The remaining five factors will receive brief discussion or none; for until these fainter patterns are independently substantiated in British social history analyses, or demonstrated in

#### FACTOR 4

##### Enlightenment

Variable No.	Loading	Description
31	.40	High government expenditure on education
32	.38	Active emigration



similar factors in other P- or R-technique analyses the labor of interpretive discussion would be premature or misleading.

In looking at the above table for Factor 4 we should note that variable 48 is also loaded there (0.66), but as typhus deaths were extremely few and sporadic (typhoid and tuberculosis being the variables we would have preferred for investigating this factor) it turns out not to be worthy of record. Although nothing like this factor appears in the U.S. study, a factor of Enlightened Affluence, tying education with prosperity, has appeared in R-technique studies with nations (11) (16), states of the union (26) and cities (26a). Affluence does not appear here, but our only marker for it (No. 34) does not in fact load *any* factor and we may wonder whether a true index of mean individual real wealth exists in our data. A measure of incidence of tuberculosis should certainly be tried in further explorations of this factor.

#### FACTOR 5

##### Slum Morale — vs — Cultural Integration

Variable No.	Loading	Description
35	42	Many families per dwelling
46	36	High frequency of homicides
47	29	High frequency of deaths from syphilis
23	27	High marriage rate
38	14	High illiteracy of women

With the particular variables of our study one might be content to label this "slum conditions". Factor 4 in the U.S. study (14) defined by "many people in gaol", "high gross death rate", "few acres per farm" and "few females per male in school attendance" has a distinct resemblance of pattern to this, though, from other associations, we called it there only "Non-conformity or Individualism" (a slum being perhaps hard to produce in so much space!). But in the R-technique national studies (11) (16) a factor of Poor Morale — vs — Cultural Integration was found quite similar to this and Hofstaetter finds a similar "slum conditions" factor in his more penetrating analysis of Thorndike's data on 300 American cities. His factor is marked by high homicide and syphilis rates and high infant death rates (plus poverty and typhoid as in our R-technique factors). (26a).

No meaning will be speculated upon for Factor 6 which follows

## FACTOR 6

Variable Loading No.		Description
32	—55	Scant emigration
42	46	High frequency of deaths from alcoholism
3	44	High average age of government ministers

Factor 7 shews one of the few substantial correlations among

## FACTOR 7

Variable Loading No.		Description
12	64	High trade union membership
30	—39	Few famous men famous in politics

factors ( $-0.47$  with Factor 2, War Stress). It has some relations to R-technique factor 6 (Scientific industriousness) which combines high trade union membership with high scientific interest and low interest in political personages.

Factor 8 has substantial loadings only on variables 35 and 20 which happen to be among the few variables (also 12 and 24) with poorly estimated communalities (resulting in their exceeding unity), wherefore it is not worth listing. Future analyses should aim specifically to bring down the communality estimates of these variables.

## 5. DISCUSSION AND SUMMARY

For those who are unfamiliar with factor analytic studies generally — and perhaps even for those who are — greater conviction that the method of analysis used is engendering more than mere “mathematical abstractions” is given by the existence of ancillary evidence of the functional unity of the factors. In psychology this has sometimes been provided by independence of the factors in experimental situations (15) and sometimes by obvious coincidence of the factors with common sense knowledge. In this case we can provide some immediate independent evidence of functional unity, as well as of meaning, by plotting a graph of the behavior of these factors over a hundred years.

Taking the more fully represented factor patterns — 1, 2, 3 and 5 — we have estimated each for each year by combining the scores of the chief variables listed above as making the given factor. The strength of factor 1 was computed by adding the

standard scores <sup>2)</sup> on variables 16, 17, 18, 21, 36, 37, 44 and 45; of Factor 2 by adding -2, 14, 19 and -22; Factor 3 by 20, 24, 26, 27 and -29; and Factor 5 by 23, 35, 46 and 47. In each case these are the highest loaded variables, all over 0.40 (with two deliberate exceptions), omitting only those which, by overlapping factors, would produce spurious correlations. The changes of these factors with time are shown in Diagram 1, where each factor is represented with approximately the same mean and standard

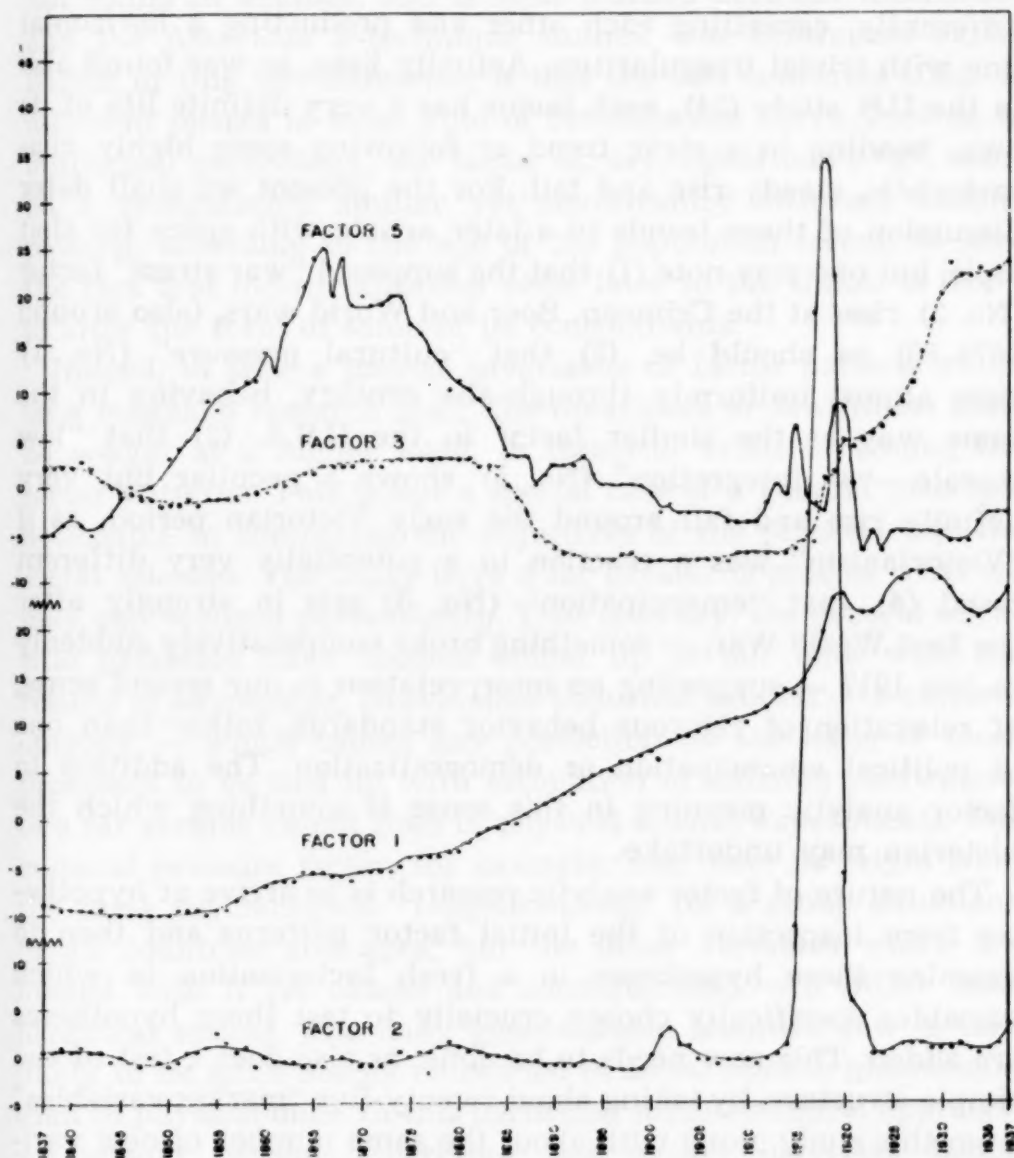


Diagram 1

<sup>2)</sup> Approximately. All variables had been brought to a twelve point scale and it was these values which were added.

deviation. (Note the scale on the left is broken into three parts, so that the absolute level of the curves is not to be considered).

It will be evident that if factor analysis simply brought together a collection of variables in each factor which, from the standpoint of a functional social psychology, is a merely chance selection, the plot of such a factor would be unlikely to show any regular trend. For the diverse variable scores would behave differently, cancelling each other and producting a horizontal line with trivial irregularities. Actually here, as was found also in the U.S. study (14), each factor has a very definite life of its own, heading in a clear trend or following some highly characteristic, steady rise and fall. For the present we shall defer discussion of these trends to a later article with space for that topic, but one may note (1) that the supposed "war stress" factor (No. 2) rises at the Crimean, Boer and World wars, (also around 1874-80) as should be, (2) that "cultural pressure" (No. 1) rises almost uniformly through the century, behaving in the same way as the similar factor in the U.S.A. (3) that "low morale -vs- integration" (No. 5) shows a peculiar but very definite rise and fall around the early Victorian period, as if "Victorianism" was a reaction to a potentially very different trend (4) that "emancipation" (No. 3) sets in strongly after the first World War, — something broke comparatively suddenly in late 1917 — suggesting an interpretation in our second sense, of relaxation of rigorous behavior standards, rather than one of political emancipation or democratization. The addition to factor analytic meaning in this sense is something which the historian may undertake.

The nature of factor analytic research is to arrive at hypotheses from inspection of the initial factor patterns and then to examine those hypotheses in a fresh factorization in which variables specifically chosen crucially to test those hypotheses are added. This now needs to be done, as also does a test of the simple structure, by taking about twenty-five "marker variables" from this study along with about the same number of new variables, to see if blind rotation leads to a similar loading pattern for the former variables. With these goals achieved the formulation of psychodynamic explanations to account for the hypothesized nature of each factor will be relatively simple.



Although a gratifying degree of similarity has now been demonstrated among the factors from independent studies — notably on the factors of Cultural Pressure, Enlightened Affluence, War Stress, Emancipation, Morale and possibly Industriousness — there also exist unquestionable differences of emphasis. This shows itself in the patterns of the common factors as well as in the existence of whole factors in one series not found in another, and is most marked between the British and the American P-technique studies. Our contingent explanation of this last difference is that the two countries stand at different phases in some kind of development curve. Just as, in individual psychology, the factor of, say, Spearman's "g", takes on a recognizably similar yet significantly different loading pattern, according to the age of the population tested, so here the fact that industrialization came later to the U.S.A. is likely to alter the form of some of its concomitants.

Indeed, to seek a precise invariance of factor pattern would be a mistaken research goal. The invariance or lawfulness must be sought at a higher level, in terms of relations among the factor patterns. This is but a special case of a general principle noticeable in comparing the objectives of the physical and the social sciences. The latter have a far greater degree of what we may call *realized particularity*. That is to say, the general scientific processes have become bound up in our data with the results of an ongoing, irreversible historical process — a cultural entropy — which causes any prediction on the basis of these processes to be tied up with estimation of existing particulars, to a far greater extent than in physical science experiments. The cultural pressure factor, for example, may take its origin from the same psychological "long-circuiting" (in a group situation) in all countries and ages; but the other variables which are loaded with it (as causes and consequences) will differ with historical setting. The truly generalizable scientific law is thus likely to lie more deeply embedded in social science phenomena than in physical ones. In this particular methodological approach we should perhaps, therefore, not expect invariant factor patterns, but only invariance of *relations* of patterns when life cycle, international situation, and the roles of other factors e.g. size, are taken into account.

To summarize:

(1) A factorization of 48 syntality, structure and population variables, measured annually on Great Britain from 1837-1937, by P-technique, has yielded effectively seven factors.

(2) Five or six of these factors have distinct resemblance to factors found either in a corresponding P-technique factorization of the U.S.A. or in an R-technique factorization of 69 countries.

(3) For discussion of the nature of these factor dimensions the reader is referred to earlier research and to future articles wholly concerned with that problem, but brief reference is made here to causes of modified loading patterns observed in the matching with other studies.

(4) Estimates made annually for four of the factors are plotted for the century, showing that they behave as distinct functional unities. However, there is somewhat more correlation (inter-action?) among these functional unities than is usually found in R-technique studies.

(5) The intensity of action (magnitude) of any variable in any year can now be expressed by the P-technique form of the usual factor analytic specification equation, as follows:

$$H_{my} = s_{m1} \cdot F_{1y} + s_{m2} \cdot F_{2y} + \dots s_{mn} \cdot F_{ny} + s_m \cdot F_{my}$$

where  $H_m$  is a particular historical manifestation (variable)  $m$  in the year  $y$ ;  $s$  is a factor loading, here called a situational index (15), the number subscript indicating the factor involved and the  $y$  indicating the year; and the  $F$ 's are the various major cultural factors found to be at work throughout the century, estimated for the given year as indicated by the subscript  $y$ .

(6) Future research is required (a) independently to check the particular simple structure adopted here (b) to enter a fresh factorization with special variables to test our hypotheses about the nature of cultural pressure, war stress, enlightened affluence, emancipation etc. and so lead to clearer psychodynamic formulations, and (c) to make P-technique factorizations of the dimensions of change in nations with different initial (R-technique) culture patterns, in order that higher order laws may emerge from observations on the similarities and dissimilarities of these patterns.

#### RESUMÉ

A Quantitative Analysis of the Changes in the Culture Pattern of Great Britain 1837-1937 by P-Technique.

Cet article est l'exposé et la démonstration d'une méthode de recherche en psychologie sociale, qui est probablement plus précise et plus puissante que d'autres méthodes employées jusque maintenant. La méthode d'analyse factorielle a été introduite récemment dans une série de recherches sur des mesures multivariées d'anthropologie culturelle et dans des expérimentations sur des groupes peu nombreux. Cette contribution continue l'application interdisciplinaire de la psychologie sur des données sociologiques et développe une nouvelle forme d'analyse factorielle nommée P-technique.

Quarante-huit variables sont choisies pour décrire les aspects psychodynamiques de la vie culturelle de la Grande Bretagne et des mesures sont faites pour chaque année de 1837 à 1937. L'intercorrélation de ces variables est alors calculée et la matrice de ces corrélations est analysée par la méthode de facteurs multiples de Thurstone. Sept facteurs ont été trouvés mais il n'y en a que cinq qui ont une saturation (loading) significative dans plus de trois variables.

L'auteur tâche d'interpréter la signification de ces changements dans le groupe social, d'abord par la comparaison avec des facteurs trouvés dans une étude analogue pour la même période publiée aux Etats-Unis. Des facteurs de pression culturelle, de tension de guerre, d'intégration et de communication améliorée semblent être actifs comme facteurs décélables dans les deux pays. Puis, comme l'existence indépendante de ces facteurs en tant que dimensions mesurables semble confirmée maintenant, un examen plus intensif de leur nature se poursuit en étudiant les «loadings patterns».

Pour pouvoir donner une interprétation plus complète, les tendances des quatre facteurs à travers le siècle étudié sont mises en graphique; on constate que les durées et les directions de leurs changements s'accordent bien avec les hypothèses concernant leur indépendance mutuelle et leur nature essentielle. Du point de vue de la psychologie de l'apprentissage et de l'étude de la personnalité, le facteur de pression culturelle a le plus grand intérêt et est interprété comme une augmentation de l'agression à cause de la frustration par suite de la complication culturelle croissante. Ce facteur se manifeste *directement* dans les hostilités internes et internationales et, *sous une forme sublimée*, dans l'augmentation de la productivité culturelle.

## BIBLIOGRAPHY

1. Abel, J. F., Illiteracy in the several countries of the world, *U. S. Bureau of Education Bulletin*, 1929, No. 4.
2. ———, *Annuaire Internationale de Statistique*, par l'Office permanent de l'Institute international de statistique, La Haye, 1916—21.
3. ———, *Annual Reports of the Registrar General, England, Wales & Scotland*. 38th, 1875, 45th, 1882, 64th, 1901, etc. London 1877, Eyne & Spottiswoode, London 1937 Arnold.
4. Benedict, R., *Patterns of Culture*, Boston. Houghton-Mifflin, 1934.
5. Brendon, J. A., *A Dictionary of British History*, London, 1877.
6. ———, *British Almanac*, Cassel & Co., limited, 1908—14, pub. from 1828—1908 by various publishers.
7. Carr-Saunders, A. M., *World Population: Past Growth & Present Trends*, Oxford: The Clarendon Press 1936.
8. Cattell, R. B., A. K. S. Cattell & R. M. Rhymer, P-technique demonstrated in determining psycho-physiological source traits in a normal individual. *Psychometrika*, 1947, 12, 267—288.
9. Cattell, R. B., Concepts and methods in the measurement of group syntality. *Psychol. Rev.*, 1948, 55, 48—63.
10. ——— and L. Wispe, The dimensions of syntality in small groups. *J. Soc. Psychol.*, 1948, 28, 57—78.
11. ———, The dimensions of culture patterns by factorization of national characters. *J. Abn. and Soc. Psychol.* 1949, 44, 443—469.
12. ———, *Personality: A Systematic Theoretical and Factual Treatment*. New York: McGraw Hill, 1950.
13. ———, The principal culture patterns discoverable in the syntal dimensions of existing nations. *J. Soc. Psychol.*, 1950, 32, 215—253.
14. ——— and M. Adelson, The dimensions of social change in the U.S.A. as determined by P-technique, *Soc. Forces*, 1952, 30, 190—201.
15. ———, *Factor Analysis, for Psychology and The Social Sciences*. New York; Harper Bros. 1952.
16. ———, An Attempt at more refined definition of the cultural dimensions of syntality in modern nations. *Amer. Soc. Rev.*, 1952, 17, 408—421.
17. ———, and G. Stice, The Psychodynamics of small Groups: An Experimental Study of Leadership, Syntality and Morale in 100 Groups. Office of Naval Research Report, 1953.
18. ———, Toward a set of standard variables for the quantitative analysis of culture patterns. (In Press)
19. ———, Hypothese on the nature of Cultural Pressure, Order Enlightenment and Morale, as cultural dimensions. (In Press)
20. Cole, G. D. H., *A Short History of the British Working Class Movement 1789—1937*. London: G. Allen & Unwin Ltd., 1937.
21. ———, *Constitutional Yearbook*, 1938; London. Harrison & Sons Ltd.
22. Durkheim, Emile, *Le Suicide*; Paris, F. Alcan, 1930.



23. ———, *Encyclopedia Britannica*: Chicago, Encyclopedia Britannica, Inc., 1949.
24. Hobhouse, L. T., G. C. Wheeler & M. Ginsberg, *The Material Culture & Social Institutions of the Simpler Peoples*. London, 1915, Chapman & Hall.
25. Hoffman, T. L., *The Mortality from Cancer throughout the world*. Newark: The Prudential Press, 1915.
26. Hofstaetter, P. R. A factorial study of cultural patterns in the U.S., *J. of Psychol.*, 1951, 32, 99—113.
- 26a. ———, "Your City" revisited: a factorial study of culture patterns. (In Press)
27. Mitchell, Wesley C., *Business Cycles*, University of California Press, Berkeley, 1913.
28. Kuczinski, R. R., *The Balance of Births & Deaths*; New York, MacMillan Co. 1928—1931.
29. Nuttall, Austin P., *The Nuttall Encyclopedia*, London & New York: F. Warne & Co., 1900, edited by the Rev. James Wood.
30. Porter, George R., *Progress of the Nation*, London, Methuen & Co. Ltd., 1912.
31. Sorokin, P. A., *Social & Cultural Dynamics*. Vol. II, New York, American Book Co., 1937.
32. ———, *Statesman's Yearbook*. 1864—1951 New York, MacMillan Co.
33. ———, *Statistical Abstracts for the United Kingdoms; Annual Abstract of Statistics*; London; Great Britain Stationary Off. London, Central Statistical Office, 1—83, 1840—1854.
34. *The Registrar-General's Statistical Review of England & Wales*. London, H. M. Stationery Office.
35. ———, *Statistical Abstract of the World*. U.S. Bureau of Statistics (Dept. of Commerce & Labor) Washington, Gov't Printing Off., 1904.
36. Thomas, W. D. & F. Znaniecki, *The Polish Peasant in Europe & America*; Chicago, Univ. of Chicago Press, 1918—20.
37. Thorndike, E. L., *Your City*. New York: Harcourt Brace. 1939.
38. Thurstone, L. L., *Multiple Factor Analysis*. Chicago. Univ. of Chicago Press, 1949.
39. Toynbee, A. J., *A Study of History*. New York: Oxford University Press, 1947.
40. Unwin, J. D., *Sex & Culture*. New York: Oxford University Press, 1934.
41. ———, *Whitaker's Almanak* 1933 London, Joseph Whitaker, 1869—1951.
42. ———, *Who's Who*. London: Black & Co., 1870—1951.
43. Williams, H. V. M., *A Determination of psychosomatic functional unities in personality by means of P-technique*. (In Press).

*Laboratoire de Psychologie de l'Hôpital Universitaire Brugmann,  
Bruxelles*

## ESSAIS D'EXPLICATION DE L'ERREUR CARACTÉRISTIQUE DANS UN MODE DE DÉLIMITATION DES RÉGIONS CUTANÉES

PAR

R. NYSSEN et J. HOZAY

A la suite de constatations faites par Jean Titeca sur des sujets hystériques, nous avons montré, dans deux travaux précédents, que si l'on sollicite un sujet normal de localiser une ligne tracée sur la peau, en fonction d'un stimulus tactile se déplaçant vers cette ligne, ce sujet présente une tendance marquée à le faire avant que le but ne soit atteint par le stimulus. En d'autres mots, il «anticipe» la localisation.

Ce phénomène d'«anticipation» est presque constant. En effet, il s'est vérifié dans la très grande majorité des 720 sujets normaux adultes chez lesquels nous l'avons exploré.

Rappelons en quelques mots la technique utilisée dans l'expérience qui a servi de base à notre recherche, c'est à dire celle dans laquelle le sujet doit localiser une ligne tracée à mi-hauteur de l'avant-bras: Trois lignes transversales sont tracées, sous les yeux du sujet, l'une au niveau du pli du coude, la seconde au pli du poignet et la troisième exactement à mi-distance des deux précédentes. Il est ensuite expliqué au sujet que le pinceau, qu'on

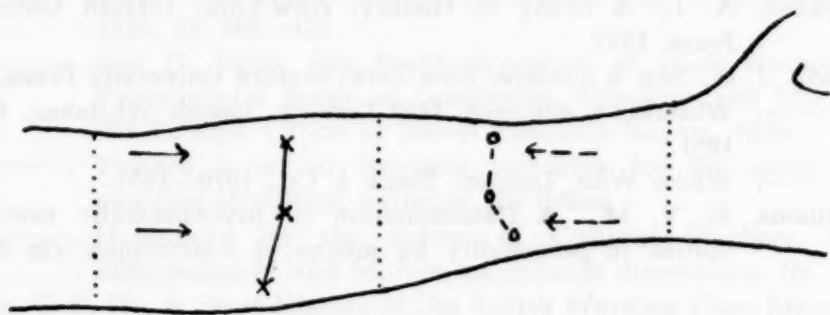


Fig. 1.

Fig. 1. Localisation d'une ligne tracée sur la face antérieure de l'avant-bras gauche.

lui montre, sera promené sur la peau, d'abord à partir du pli du coude vers la ligne qui sépare la moitié supérieure de la moitié inférieure de l'avant-bras, puis à partir du poignet vers cette même ligne. La consigne lui est donnée, qu'ayant les yeux bandés, il devra dire «stop» dès qu'il estimera que le pinceau aura atteint la ligne du milieu.

Le pinceau effectue successivement trois passages descendants, puis trois passages ascendants. Le premier passage descendant suit la ligne médiane, le second le bord externe et le troisième le bord interne de la face palmaire de l'avant-bras; les passages ascendants sont effectués dans le même ordre. Chaque passage du pinceau se fait avec une même régularité de pression (environ 7 grammes) et de vitesse ( $\frac{1}{2}$  cm par sec.) jusqu'au moment du signal du sujet. Chaque localisation de la ligne faite par le sujet est marquée sur la peau, par l'un des expérimentateurs, à l'aide d'un mince tampon d'ouate imbibé de colorant. Après l'expérience, la ligne transversale et les six points de localisation sont calqués, ce qui permet de mesurer en millimètres l'erreur de chaque localisation et de comparer les résultats entre eux.

Cette technique fut appliquée dans tous les groupes d'expériences où il s'agissait de localiser une ligne transversale. Pour les groupes où le sujet avait comme tâche de localiser une ligne longitudinale, les passages médians, externes et internes étaient remplacés par des passages supérieurs, médians et inférieurs. Nous avons examiné le phénomène à des régions très différentes: à mi-hauteur des avant-bras, à mi-hauteur des jambes, à plusieurs travers de doigt au-dessus ou au-dessous du pli du coude, à plusieurs travers de doigt au dessus du pli du poignet, au pli du coude et au pli du poignet, au milieu du dos et de la poitrine, au milieu du front et du nez.

Aussi bien dans les délimitations longitudinales que dans les délimitations transversales l'«anticipation» ne cesse de se produire avec une fréquence et une importance très marquées. Elle mesure habituellement plusieurs centimètres; dans la localisation d'une ligne transversale à mi-hauteur de la jambe l'«anticipation» atteint même une moyenne de 71 mms.

Les «transgressions» de la limite par le stimulus ne se produisent qu'exceptionnellement. Elles sont en général nettement beaucoup plus petites que les «anticipations».

Enfin, ce phénomène de l'«anticipation» se rencontre avec

une importance et une fréquence égales chez les intellectuels et les sujets non cultivés, chez les sujets intelligents et chez les moins doués.

Cette localisation indirecte d'une ligne, en fonction du stimulus tactile se dirigeant vers elle, est plus complexe que la localisation directe qui consiste en la désignation, par le sujet, de l'endroit cutané touché. Celle-là diffère de celle-ci:

- a) par la tendance prononcée à «anticiper» la localisation,
- b) par une plus grande importance de l'erreur,
- c) par un plus grand sentiment d'incertitude ou d'incapacité du sujet de la localisation.

Nous voudrions examiner, dans le présent travail, quels peuvent être le ou les facteurs de cette erreur spécifique de délimitation ou de localisation.

Parmi les facteurs qui nous paraissent devoir être envisagés, il y a lieu de distinguer des facteurs déterminants et des facteurs adjuvants.

Nous considérons comme facteurs déterminants ceux qui paraissent pouvoir exercer une action efficiente sur la production même du phénomène. Par facteurs adjuvants nous entendons ceux qui sont susceptibles d'exercer une action favorisante sur la grandeur de l'«anticipation». On peut cependant admettre que l'action favorisante puisse constituer, dans certaines proportions, aussi une action déterminante de la production même de l'anticipation. En faveur de cette conception plaide entre autres le fait, que nous avons remarqué, dans nos recherches, que dans chaque groupe d'expériences il existe une concordance positive entre la grandeur moyenne des anticipations et le degré de fréquence du phénomène.

#### FACTEURS DÉTERMINANTS

Nous paraissent susceptibles d'être envisagés comme facteurs déterminants:

- 1) la tendance à un mode de localisation,
- 2) l'influence de certains repères sur le mode de localisation,
- 3) le mode de représentation des dimensions corporelles,



- 4) l'appréciation du temps de passage du stimulus nécessaire pour atteindre la ligne,
- 5) le mode d'estimation de la longueur du déplacement du stimulus tactile.

### 1. *La tendance à un mode de localisation*

Barth et aussi Pillsbury ont constaté, dans des recherches faites sur le bras, que la localisation directe (Ortsinn) présentait une tendance constante à se faire vers l'extrémité distale du membre.

Cette constatation est contredite par celles de Volkmann, de Koltenkamp et Ulrich, de Haines, de Mc. Dougall.

Nous-mêmes avons constaté que la localisation directe (par l'index) de la ligne qui sépare la moitié supérieure de la moitié inférieure de l'avant-bras, étudiée chez 130 sujets, donnait les résultats suivants: sur 390 épreuves, 176 localisations ont été faites au-dessus de la ligne avec un écart moyen de 19,9 mms, 64 sur la ligne même, et 150 au-dessous de la ligne, avec un écart moyen de 18,2 mms. Ces nouvelles données ne concordent donc pas avec les conclusions de Barth et de Pillsbury.

Au demeurant, s'il existait réellement une tendance plus ou moins générale à localiser vers l'extrémité du membre, l'intervention de cette tendance serait aussi en contradiction avec le fait que les anticipations sont du même ordre dans la moitié proximale et la moitié distale des segments de membres explorés.

Signalons en outre que les «anticipations» lors des mouvements transversaux et symétriques du stimulus tactile ne sauraient trouver leur explication dans une telle tendance.

### 2. *L'influence de certains repères sur le mode de localisation*

V. Henri a insisté sur l'importance des points de repère, tels que certains points saillants de la peau, les bords, les plis, etc., auxquels les sujets rapportent le point touché. On remarque, aux dires de cet auteur, que presque toujours l'erreur de localisation est commise dans la direction des points de repère que le sujet a employés pour localiser le contact. Mais il ajoute que, quand la distance entre le point touché et le pli est grande, la distance est souvent exagérée par le sujet.

Beaucoup de nos sujets nous ont déclaré qu'ils tâchaient de rapporter la ligne qu'ils devaient localiser à des points de repère

et en particulier à des plis de la peau ou à une articulation plus ou moins voisine. Très souvent nous avons remarqué que des sujets essayaient de faciliter la localisation en exécutant de petits mouvements de flexion au niveau du coude ou du poignet.

On peut donc envisager l'intervention de ces repères dans le phénomène de l'«anticipation». Ce facteur ne paraît cependant pouvoir intervenir qu'en ordre très subsidiaire. En effet:

- a) l'«anticipation» est encore bien marquée, quand précisément il s'agit de localiser un pli important tel celui du coude ou celui du poignet;
- b) quand il s'agit de localiser une ligne tracée à trois travers de doigt du pli du coude ou de pli du poignet, le sujet, au lieu d'«anticiper» d'avantage quand le stimulus se meut du pli vers la ligne que quand il se meut en sens inverse, anticipe d'une façon semblable dans les deux sens, et souvent même anticipe le plus quand le stimulus se dirige vers le pli.

### 3. *Le mode de représentation d'une distance cutanée*

Un autre facteur éventuel qui nous paraît devoir être envisagé est le mode de représentation d'une distance entre deux points cutanés.

D'après Wundt, si l'on demande à un sujet de reproduire par représentation visuelle la distance qu'il perçoit entre les deux stimulus tactiles d'un esthésiomètre, la distance reproduite est considérablement plus petite que l'écart entre les pointes esthésiométriques. Wundt met ce fait en relation avec la réduction de l'image corporelle par rapport aux dimensions réelles de notre corps et de ses parties.

Nous avons vérifié l'intervention éventuelle du fait relaté par Wundt, dans le phénomène d'«anticipation», en faisant appel à la reproduction immédiate d'une distance esthésiométrique actuellement perçue.

Sous le contrôle de sa vue, le sujet devait indiquer, de l'index et du pouce, sur le bord de la table d'examen, l'écart perçu entre les pointes en ivoire d'un compas, appliquées sur l'avant-bras. Les distances perçues à l'avant-bras gauche devaient être indiquées par la main droite et celles perçues à l'avant-bras droit par la main gauche. L'avant-bras reposait confortablement sur un coussin et était caché aux yeux du sujet par un écran. Chaque sujet devait reproduire successivement

12 applications esthésiométriques. Toutes les applications se faisaient dans un ordre bien déterminé: chez 30 sujets on commençait par l'application d'une distance de dix cms, suivie d'une distance de 8 cms, puis de six cms. La pointe proximale de ces trois distances était appliquée chaque fois au niveau du coude gauche. Les trois applications suivantes, également de dix, huit et six cms, se faisaient aussi dans un ordre décroissant; la pointe distale était cette fois posée sur le poignet gauche. Chez les mêmes sujets la recherche était reprise à l'avant-bras droit, par trois applications à partir du coude et trois applications à partir du poignet. Ces applications se faisaient avec les mêmes distances esthésiométriques que celles employées pour l'avant-bras gauche, mais cette fois dans un ordre croissant: six, huit et dix cms. Comme il y avait lieu de tenir compte de la possibilité d'une influence de contraste, exercée par l'application précédente sur l'appréciation de la distance perçue, — influence de contraste signalée déjà par Wundt — un second groupe de 30 sujets a subi l'expérience par des applications faites dans un ordre inverse.

Ces recherches, effectuées donc chez 60 sujets différents, ont confirmé, dans certaines limites, la règle d'estimation des distances esthésiométriques formulée par Wundt:

- a) Un nombre important de ces 60 sujets (environ 50 %) ont déclaré qu'ils n'avaient pas l'impression de s'être représenté visuellement l'écart perçu. Les autres avaient plus ou moins nettement l'impression d'avoir eu recours à une représentation visuelle, soit de l'écart appliqué sur une surface indéterminée, soit de l'écart des deux pointes sur l'avant-bras.
- b) Chacun des 60 sujets a subi, aux coudes, deux applications de 10 cms, autant d'applications de 8 cms et autant de 6 cms. La grande majorité de ces sujets a reproduit les distances en les réduisant.

Il est intéressant de constater que le nombre de réductions diminue avec la distance appliquée. En effet, sur les 120 applications de 10 cms de distance, nous notons 3,4 fois autant de réductions que d'augmentations; sur les 120 applications de 8 cms, il y a 2,8 fois autant de réductions que d'augmentations; enfin, sur les 120 applications de 6 cms, il n'y a plus que 1,9 fois autant de réductions que d'augmentations.

Une constatation semblable a été faite aux poignets: 2,1 fois

autant de réductions que d'augmentations sur les 120 reproductions des écarts de 10 cms; 1,65 fois autant de réductions que d'augmentations sur les 120 écarts de 8 cms; enfin 1,2 fois autant de réductions que d'augmentations des 120 écarts de 6 cms.

Le fait relaté par Wundt se confirme donc par ces constatations. Il réclame toutefois des précisions, compte tenu de nos conditions expérimentales:

- a) le nombre et l'importance des réductions diminuent avec les distances que le sujet doit reproduire;
- b) le nombre et l'importance des réductions varient avec la région sur laquelle les écarts esthésiométriques sont appliqués.

Il va sans dire que ces constatations n'impliquent pas par elles-mêmes que la tendance des sujets à sous-estimer une distance entre deux points cutanés serait une cause efficiente du phénomène de l'anticipation.

Il existe au contraire des motifs sérieux pour croire qu'il n'y a pas de relation causale entre le mode d'estimation d'une telle distance et le phénomène de l'«anticipation»:

- a) Dans les expériences où le sujet devait localiser une transversale tracée à mi-hauteur de l'avant-bras, l'importance des «anticipations» aux passages du stimulus du poignet vers la ligne était de même ordre que celle des anticipations aux passages descendants. Or, comme les réductions de reproduction des applications esthésiométriques sont bien moins importantes pour la moitié inférieure que pour la moitié supérieure de l'avant-bras, les «anticipations» ascendantes auraient dû être bien inférieures aux «anticipations» descendantes si le fait de l'«anticipation» était dû en tout ou en grande partie à la réduction de l'image corporelle d'après Wundt, ou à la tendance à la réduction de la reproduction d'une distance cutanée.
- b) Si la réduction dans l'estimation d'une distance entre deux points était à retenir comme cause efficiente de l'anticipation, le groupe des sujets qui ont présenté la réduction de reproduction la plus importante aurait vraisemblablement présenté les «anticipations» les plus élevées. Par contre, les sujets qui ont réduit le moins les écarts esthésiométriques ou même les ont augmentés, auraient vraisemblablement présenté les «anticipations» les plus petites ou auraient transgressé dans la localisation de la ligne. Or, que constatons-nous?



- 1) Parmi les 15 sujets qui, sur les 60, ont réduit le plus l'écart de 10 cms appliqué au coude gauche (les écarts reproduits par eux varient de 2,25 à 6,5 cms), 7 seulement dépassent la moyenne des anticipations;

parmi les 15 sujets qui ont réduit le moins l'écart de 10 cms appliqué au coude gauche ou l'ont augmenté (les écarts reproduits par eux varient de 9,25 à 15 cms) 6 seulement donnent des anticipations inférieures à la moyenne.

- 2) Parmi les 15 sujets qui ont réduit le plus l'écart de 10 cms appliqué au coude droit (les écarts reproduits varient de 3,5 à 6,5 cms), 5 seulement dépassent la moyenne des anticipations;

parmi les 15 sujets qui ont réduit le moins l'écart de 10 cms appliqué au coude droit ou l'ont augmenté (les écarts reproduits varient de 8,5 à 17 cms), 7 seulement donnent des anticipations inférieures à la moyenne.

- 3) Parmi les 15 sujets qui ont réduit le plus l'écart de 10 cms appliqué au poignet gauche (les écarts reproduits varient entre 4 et 7,5 cms), 7 seulement dépassent la moyenne des anticipations;

parmi les 15 sujets qui ont augmenté le plus l'écart de 10 cms appliqué au poignet gauche (les écarts reproduits varient entre 10,5 et 15,5 cms), 4 seulement ont fourni des anticipations inférieures à la moyenne.

- 4) Enfin, des 15 sujets qui ont réduit le plus l'écart de 10 cms, appliqué au poignet droit (écarts reproduits variant entre 4,75 et 7,5 cms), 6 seulement anticipent au-dessus de la moyenne;

parmi les 15 sujets qui ont augmenté le plus l'écart de 10 cms appliqué au poignet droit (écarts reproduits varient entre 10 et 14,5 cms), 6 seulement ont donné des anticipations inférieures à la moyenne.

La tendance à réduire, dans sa représentation, une distance cutanée ne paraît donc pas intervenir d'une façon efficiente dans la production du phénomène de l'anticipation.

#### 4. *L'appréciation du temps de passage du stimulus nécessaire pour atteindre la ligne*

L'étude de ce facteur nous paraît très complexe, car d'une part le sujet n'a pas d'expérience préalable de la durée en question et d'autre part l'appréciation de cette durée sera

fonction de la vitesse de passage du stimulus et de la distance que le stimulus doit parcourir pour atteindre la ligne.

La méthode qui nous paraît encore pouvoir le mieux convenir pour l'exploration de l'appréciation de la durée d'un stimulus tactile, indépendamment de la vitesse et de la longueur de son passage, est une méthode indirecte, c'est-à-dire celle basée sur l'appréciation d'une durée déterminée d'un stimulus tactile, appliqué sur un endroit fixe. Malheureusement, par l'absence de déplacement du stimulus sur la peau, et, non moins, par la nécessité de fixer l'attention du sujet sur la durée de l'attouchement, cette exploration n'a plus qu'un rapport réduit avec les conditions psychologiques dans lesquelles se produit le phénomène de l'«anticipation».

Pour pouvoir examiner l'intervention éventuelle de l'estimation de la durée du stimulus dans le phénomène de l'«anticipation» nous avons pensé qu'il était indiqué de rechercher s'il existe une corrélation entre l'estimation de cette durée d'une part, et le degré de l'anticipation d'autre part.

Nous avons opéré sur cent sujets. Chez chacun d'eux nous avons exploré leur façon d'apprécier la durée d'un stimulus tactile produit par la rotation de la touffe d'un pinceau fin au niveau de la mi-hauteur de la face palmaire de l'avant-bras gauche. La durée de ce stimulus local et continu était de 20 secondes c.à.d. une durée correspondant approximativement à celle du passage linéaire du pinceau dans la plupart de nos expériences faites pour l'étude du phénomène de l'«anticipation». Chez chaque sujet ce même stimulus, d'une même durée, a été appliqué trois fois avec un intervalle de 15 secondes. Le sujet avait reçu comme consigne de faire attention à la durée de chaque application tactile, sans toutefois compter en lui-même et sans tâcher d'évaluer le temps en secondes. Il avait été dit au sujet que la durée était la même à chacune des trois applications.

Après cette initiation, le même stimulus tactile continu était appliqué au même endroit de l'avant-bras, mais cette fois le sujet avait à dire «stop», quand il estimait que l'attouchement avait atteint la même durée que celle des attouchements précédents. Cette appréciation de durée était soigneusement notée au cinquième de seconde. Pendant toute la durée de l'expérience le sujet avait les yeux couverts d'un bandeau.

De ces cent sujets, 59 ont arrêté la quatrième application du

stimulus après une durée inférieure à 20 secondes, soit après une durée moyenne de 15,86 secondes; 6 sujets l'ont évaluée d'une façon exacte, tandis que 35 seulement ont dépassé le temps donné: 23,94 secondes en moyenne. Ces constatations tendent à établir l'existence d'une certaine tendance à la réduction dans l'estimation de la durée du stimulus.

Chez chacun des cent sujets, immédiatement après la recherche de l'estimation de la durée du stimulus local, nous avons examiné, en appliquant la technique habituelle, la façon dont ils localisaient une ligne transversale à mi-hauteur de la face palmaire de l'avant-bras gauche, en fonction d'un stimulus tactile se dirigeant vers cette ligne à partir du coude et aussi à partir du poignet.

Les corrélations (Bravais-Pearson) ont été calculées entre les évaluations de la durée du stimulus s'exerçant sur place d'une part, et la longueur de chacun des 6 passages du pinceau d'autre part. Ces corrélations ont été trouvées nulles ou insignifiantes:  $-.02, -.01, .08, .25, .10, .10$ .

Cette absence de corrélation se trouve confirmée par le fait suivant: les 25 sujets donnant l'estimation de la durée la plus longue n'ont pas fourni en moyenne des longueurs de passages sensiblement supérieures à celles fournies par les sujets qui ont estimé les durées les plus courtes. En effet, ces moyennes sont de 8,97 cms pour le premier et de 8,76 cms pour le second groupe.

Cette absence de relation entre la façon d'apprécier la durée d'un stimulus tactile et celle de localiser la ligne en fonction du stimulus se déplaçant vers elle nous prive d'un argument sérieux en faveur de l'intervention de l'appréciation de la durée du stimulus mobile dans le phénomène de l'«anticipation».

##### 5. *Le mode d'estimation de la longueur du déplacement du stimulus tactile*

Il ne nous a pas paru sans intérêt d'examiner si l'estimation de l'amplitude du mouvement d'un stimulus tactile se déplaçant en ligne droite et continue, avec une vitesse régulière, sur la peau, ne se fait pas aussi avec un certain degré d'«anticipation» dans la réponse du sujet. En d'autres mots, l'«anticipation» de la réponse se produit-elle encore quand le sujet n'a plus à s'occuper de la localisation d'une ligne limite mais simplement de l'estimation de la longueur du trajet du stimulus?



Par l'expérience décrite ci-après, et faite chez 100 sujets, l'existence évidente de la tendance à «anticiper» dans l'estimation de la longueur du trajet du stimulus tactile a pu être établie. Chacun de ces sujets devait tenir entre l'index et le pouce droit une réglette non graduée d'une longueur de 12 cms. Pendant la durée de l'expérience le sujet devait regarder cette réglette. Des passages du pinceau étaient effectués sur la face antérieure de l'avant-bras gauche. Les trois premiers passages étaient effectués successivement à partir de 5 cms en-dessous du pli du coude et étaient dirigés longitudinalement vers le poignet. Le premier passage était fait le long de la ligne médiane; le deuxième le long du bord externe et le troisième le long du bord cubital de l'avant-bras. La vitesse était celle utilisée dans la grande majorité de nos expériences soit d'environ  $\frac{1}{2}$  cm. par seconde. Trois passages semblables étaient dirigés vers le coude, à partir de 5 cms au-dessus du pli du poignet. Le sujet avait comme consigne d'avertir dès que le pinceau aurait parcouru un trajet de la longueur de la réglette tenue dans la main droite. Les passages du stimulus étaient effectués sous un écran, afin de les soustraire à la vue du sujet.

Cette expérience nous a donné les résultats suivants: sur les 300 trajets descendants, 72,7 % ont été réduits par les sujets. La réduction moyenne était de 36,3 mms. Des 100 sujets, 57 ont raccourci le trajet aux trois passages. 76,3 % des 300 trajets ascendants ont été réduits d'une moyenne de 36,2 mms. Des 100 sujets, 64 ont raccourci les trois passages.

Pour étudier s'il y a relation entre cette tendance manifeste à l'«anticipation» dans l'estimation de la longueur du trajet du stimulus d'une part, et l'«anticipation» dans la localisation de la ligne d'autre part, nous avons:

- 1) appliqué chez les mêmes 100 sujets l'expérience de la localisation de la ligne transversale tracée à mi-chemin entre le coude et le poignet gauche, en fonction du stimulus tactile se déplaçant vers cette ligne à partir du coude, et à partir du poignet;
- 2) recherché les coefficients de corrélation (Bravais-Pearson) entre les résultats obtenus dans les deux expériences appliquées à ces 100 sujets.

La recherche de ce coefficient de corrélation se justifiait:

- a) par l'existence d'une façon *individuelle*, tout au moins au cours



de la séance d'examen, d'estimer la longueur du trajet du stimulus tactile aux différents passages et,

- b) par l'existence d'une façon individuelle de localiser une ligne-limite.

En effet, l'existence de ces modes individuels est établie par des corrélations élevées entre les longueurs des différents passages à chacune des expériences:

- a) pour l'expérience de l'estimation du déplacement du stimulus, nous avons trouvé entre les trois passages descendants: .67, .75, .65, et entre les trois passages ascendants: .83, .83, .78;  
b) pour l'expérience de la localisation nous avons obtenu entre les trois passages descendants: .85, .86, .79, et entre les trois passages ascendants: .78, .83, .79.

Le caractère individuel de ces deux actes étant établi, nous avons recherché la relation quantitative entre eux, en établissant le coefficient de corrélation entre chacun des passages de la première expérience et le passage homonyme de la deuxième expérience.

Ce calcul nous a donné:

pour les passages descendants	médians:	.22, $\pm$ 0,064,
" " "	externes:	.30, $\pm$ 0,061,
" " "	internes:	.37, $\pm$ 0,058,
" " " ascendants	médians:	.38, $\pm$ 0,057,
" " "	externes:	.39, $\pm$ 0,057,
" " "	internes:	.43, $\pm$ 0,055.

Ces coefficients sont faibles, mais n'en sont pas moins indicatifs d'une relation entre les deux modes d'«anticipation».

L'existence de cette relation se trouve d'ailleurs confirmée par le fait que, pour *chacun* des 6 passages, le groupe des 50 sujets ayant présenté les trajets les plus longs dans l'expérience de la localisation est aussi celui qui fournit une moyenne sensiblement plus élevée que la moyenne obtenue chez les 100 sujets, au passage homonyme de la seconde expérience. Cette supériorité à la moyenne était pour les 6 passages respectivement de 10 %, 7 %, 8 %, 9 %, 10 %, 11 %.

Il va de soi que, dans ces conditions, le groupe des 50 sujets ayant fourni les trajets les plus courts dans la première expérience est aussi celui qui a fourni une moyenne sensiblement inférieure à la moyenne de la seconde expérience.

Il y a donc lieu d'admettre une relation entre la façon de localiser une limite linéaire en fonction d'un stimulus tactile qui se déplace vers cette ligne d'une part et celle d'estimer la longueur d'un trajet d'un stimulus semblable d'autre part. Il va cependant de soi que la seule constance de cette relation n'établit pas l'existence d'un lien de cause à effet entre ces deux comportements. Nous y reviendrons, dans l'interprétation des résultats de ce travail.

#### FACTEURS ADJUVANTS

Comme facteurs adjuvants, nous pensons devoir envisager:

- 1) la nature du stimulus;
- 2) la vitesse du déplacement du stimulus;
- 3) la longueur du trajet de ce stimulus;
- 4) les degrés de schématisation de la ligne à localiser.

##### 1. *La nature du stimulus:*

Nous avons décrit, dans notre travail précédent,<sup>1)</sup> les expériences avec utilisation d'un pinceau fin et celles où ce pinceau était remplacé par une roulette dentée, ou une série de piqûres légères ou une tige en bois à pointe mousse. Nous y avons constaté une certaine tendance à l'accroissement de l'«anticipation», et dans sa grandeur et dans sa fréquence, si on remplaçait le stimulus tactile faible par des excitations cutanées plus intenses et plus profondes. Il ne nous paraît pas nécessaire d'y revenir.

##### 2. *La vitesse du déplacement du stimulus:*

Il est facile de constater, même sans recherche méthodique, que le degré de l'«anticipation» a tendance à varier en sens inverse de la vitesse du déplacement du stimulus. Nous pouvons vérifier cette constatation en mesurant les anticipations, à une vitesse lente et à une vitesse rapide déterminée du stimulus. A l'aide d'un chariot mû par une vis sans fin dont la rotation était entretenue par un mécanisme d'horlogerie avec régulateur de Foucault, nous avons déplacé mécaniquement le pinceau à une vitesse constante de 0,48 et de 1,42 cm. par seconde. Les sujets devaient localiser une ligne transversale tracée à mi-hauteur sur la face palmaire de l'avant-bras droit. Cette loca-

<sup>1)</sup> Acta Psychologica, 8: 2, 1951.

lisation devait se faire, comme dans les expériences 1 et 2 de notre article précédent, en fonction de déplacements du stimulus vers la ligne, d'abord à partir du coude, ensuite à partir du poignet. Chez 15 sujets la vitesse lente était utilisée avant la vitesse rapide, et chez 15 autres la vitesse rapide précédait la vitesse lente, ce qui nous a permis de constater que l'ordre de l'application des vitesses n'avait pas d'influence sur l'amplitude des anticipations. Cette recherche nous a appris, qu'en remplaçant la vitesse lente par la rapide, nous avons réduit de 37,8 % l'amplitude moyenne des anticipations aux passages descendants et de 31,4 % celle aux passages ascendants. Cette diminution nette de l'amplitude de l'anticipation par l'augmentation de la vitesse du stimulus se comprend déjà aisément quand on songe que le signal du sujet n'est donné qu'après une hésitation plus ou moins longue, que ce signal représente un temps de réaction suivi d'un second temps de réaction constitué par l'arrêt du stimulus par l'expérimentateur. Pendant que ces trois temps s'additionnent, le stimulus parcourt une étendue proportionnelle à sa vitesse. Mais, il y a lieu de penser qu'un ou plusieurs autres facteurs encore interviennent dans l'influence négative de la vitesse du stimulus sur l'amplitude de l'anticipation. Ainsi Vierordt avait déjà noté qu'un mouvement continu d'une pointe sur la peau paraît avoir une amplitude d'autant plus faible que la vitesse est grande. S'il en est ainsi, le sujet doit, pour une vitesse de stimulus plus élevée avoir tendance à laisser parcourir par ce stimulus une étendue plus grande.

### 3. *L'étendue parcourue par le stimulus:*

L'intervention de ce facteur est à considérer comme complexe. En effet, nous devons envisager:

- a) l'influence possible de la *distance* du point de départ du stimulus par rapport à la ligne à localiser;
- b) la *durée* d'application du stimulus variant inévitablement avec l'étendue;
- c) le mode d'estimation (sus- ou sous-estimation) de l'étendue à parcourir par le stimulus; en effet, ce mode d'estimation peut être fonction de la grandeur de l'étendue, vue par le sujet avant le bandage des yeux.

Il nous paraît toutefois impossible de dissocier ces composantes dans l'examen expérimental du facteur qui nous occupe.

Les conditions expérimentales de cet examen sont analogues à celles décrites plus haut, avec cette différence toutefois que chez 15 sujets, dans une première séance, les passages se faisaient à partir de 12 cms. de distance de la ligne à localiser et dans une seconde séance, à partir de 6 cms de cette ligne. Chez 15 autres sujets, les passages partaient dans une première séance, à 6 cms, et dans une seconde séance, à 12 cms. de la ligne. Cette inversion de l'ordre des épreuves a été faite afin de neutraliser autant que possible un transfert éventuel des résultats d'une épreuve à l'autre. Nous n'avons cependant pas relevé l'existence d'un tel transfert, en comparant entre-eux les résultats obtenus dans chacun des deux groupes.

Alors que, aux 90 passages descendants, nous obtenons à la grande distance 83 % d'«anticipations», nous n'en retrouvons à la petite distance que 74 %. Pour les 90 passages ascendants nous trouvons aux deux distances les mêmes pourcentage d'«anticipations», soit 87 %.

La valeur moyenne des «anticipations» aux passages descendants est de 38,3 mms. quand le pinceau part à 12 cms. de la ligne, et de 20,5 mms. seulement dans l'autre expérience. Aux stimulus ascendants, la moyenne des «anticipations» est de 32,4 mms, aux passages longs, et de 25,9 mms aux passages courts.

Ces constatations tendent à établir que l'étendue parcourue par le stimulus intervient positivement dans l'importance de l'«anticipation».

#### 4. *Le degré de schématisation de la ligne à localiser:*

Nous avons réalisé trois expériences différentes dans lesquelles les possibilités de schématisation de la ligne limite allaient en diminuant de la première à la troisième expérience.

##### *1ère expérience:*

Une ligne transversale est tracée à l'encre le long du pli du coude droit et une ligne semblable le long du pli du poignet droit; une troisième ligne transversale est tracée sur la face palmaire de l'avant-bras, à mi-hauteur entre le coude et le poignet. L'attention du sujet est attirée sur la situation exacte de cette ligne. Puis, *sous ses yeux*, le pinceau effectue successivement les trois passages habituels, du coude jusqu'à la ligne transversale médiane. Après cette initiation, les yeux du sujet sont bandés, et les trois passages sont répétés avec consigne



d'indiquer le moment où le sujet croit que le pinceau atteint la ligne limite.

Après avoir enregistré les résultats, nous opérons d'une façon semblable en pratiquant, après initiation préliminaire, les trois passages du poignet vers la ligne limite.

Cette expérience a été faite chez 60 sujets.

Sur les 180 passages descendants, 70,5 % donnent lieu à une localisation «anticipée». La moyenne de ces «anticipations» est de 28,3 mms. ( $\sigma = 16,2$ ;  $v = 57,25$ ).

Sur les 180 passages du poignet vers la ligne limite, il y a 94,4 % d'«anticipations» d'une moyenne de 30,4 mms. ( $\sigma = 16,8$ ;  $v = 55,03$ ).

### *2ème expérience:*

Les conditions sont celles décrites dans l'introduction de ce travail; ce sont donc les mêmes que celles de l'expérience précédente, mais sans initiation préliminaire au stimulus.

60 sujets ont servi à cette expérience.

Sur les 180 passages descendants, nous enregistrons 88 % d'«anticipations». Moyenne: 26,4 mms. ( $\sigma = 14,5$ ;  $v = 54,92$ ).

Sur les 180 passages ascendants, il y a 95,5 % d'«anticipations». Moyenne: 31,4 mms. ( $\sigma = 17,7$ ;  $v = 56,37$ ).

### *3ème expérience:*

Le sujet a les yeux bandés dès le début de l'expérience. On trace les trois lignes transversales sur l'avant-bras droit, tout en expliquant au sujet le niveau auquel on les effectue. Puis sont pratiqués les passages du pinceau descendants et ascendants comme dans les deux expériences précédentes. Comme dans celles-ci, le sujet est invité à localiser la ligne limite en fonction du stimulus qui se déplace vers elle.

Cette expérience a été appliquée chez 60 sujets.

Sur les 180 passages descendants, 99 % donnent une «anticipation». Moyenne de 52,8 mms. ( $\sigma = 20,2$ ;  $v = 32,02$ ).

Sur les 180 passages ascendants, il y a 96 % d'«anticipations». Moyenne: 46,9 mms. ( $\sigma = 20,6$ ;  $v = 43,92$ ).

Ces résultats nous indiquent que l'importance des «anticipations», tant aux passages descendants qu'aux ascendants, reste pratiquement la même si nous augmentons les possibilités de

schématisation des passages du stimulus par voies visuelle et tactile.

Par contre, la différence entre les résultats des 2ème et 3ème expériences indique toute l'importance de la schématisation visuelle de la *ligne à localiser*. Cette schématisation visuelle réduit sensiblement la grandeur de l'«anticipation», tout au moins au niveau du bras.

Nous avons tenu à vérifier ces constatations relatives à l'influence de la schématisation du stimulus mobile et de la ligne à localiser en reprenant à la jambe, des expériences semblables chez de nouveaux sujets, répartis en trois groupes de 60. La ligne à localiser est tracée exactement à mi-chemin entre une première ligne transversale passant au niveau de la pointe de la rotule et une seconde au niveau des malléoles. A chaque application de l'expérience, l'attention du sujet est attirée sur la situation exacte de la ligne à localiser.

Chez les 60 sujets initiés d'après la méthode suivie dans la première expérience sur le bras, nous relevons, à la descente, 95,6 % d'anticipations. Moyenne: 53,1 mms. ( $\sigma = 27,56$ ;  $v = 51,90$ ). Sur les 180 passages ascendants, il y a 91,7 % d'anticipations. Moyenne: 47,5 mms. ( $\sigma = 29,88$ ;  $v = 62,90$ ).

A la deuxième expérience, nous obtenons aux passages descendants 96,1 % d'anticipations. Moyenne: 56,2 mms ( $\sigma = 27,06$ ;  $v = 48,15$ ). Sur les 180 passages ascendants, nous trouvons 89,4 % d'anticipations. Moyenne: 49,6 mms. ( $\sigma = 27,75$ ;  $v = 55,95$ ).

Enfin, dans l'expérience sans présentation visuelle de la ligne à localiser, nous enregistrons 99 % d'anticipations. Moyenne: 62,1 mms. ( $\sigma = 26,96$ ;  $v = 43,42$ ). Les 180 passages ascendants nous donnent 95,6 % d'anticipations. Moyenne: 58,4 mms. ( $\sigma = 28,18$ ;  $v = 48,25$ ).

Cette vérification expérimentale sur la jambe nous donne une différence des résultats plus sensible entre les deux premières expériences que celle obtenue au bras. Par contre, les différences entre les anticipations de la 3ème expérience et celles de la seconde sont moins prononcées que dans les expériences équivalentes au bras. Toutefois, elles sont encore nettement supérieures aux différences entre les anticipations de la première et celles de la seconde expérience.

Nous pensons pouvoir conclure de ces recherches faites au bras et à la jambe: 1°) qu'il subsiste un certain doute quant à

l'influence du degré de schématisation du parcours du stimulus sur l'importance de l'«anticipation».

2°) que la visualisation préliminaire de la ligne à localiser réduit nettement la grandeur de l'anticipation.

#### INTERPRÉTATION DES RÉSULTATS

L'étude expérimentale des facteurs considérés comme *adjuvants* nous a fourni des données positives quant à leur action. Il va de soi cependant que cette action ne saurait avoir de valeur explicative suffisante pour ce qui concerne la production même du phénomène de l'«anticipation».

Quant aux facteurs *déterminants* envisagés, seul le mode d'estimation de la longueur du déplacement du stimulus tactile se montre en relation avec le phénomène de l'«anticipation» dans la localisation d'une limite linéaire, en fonction d'un stimulus tactile qui se déplace vers elle. Il n'est cependant pas permis d'attribuer à cette relation, par sa seule constatation, une signification de cause à effet. Cette relation peut tout aussi bien trouver son explication dans l'existence d'un facteur commun à l'«anticipation» relevée dans les deux expériences. Nous ne voyons pas jusqu'à présent comment nous pourrions départager ces deux possibilités.

Dans le même ordre d'idées, nous vient à l'esprit la question de savoir si le phénomène de l'«anticipation», étudié au cours de ce travail, ne se rattache pas à un ensemble de manifestations semblables qui auraient un facteur déterminant commun. Parmi ces manifestations, on peut citer, à côté du phénomène de l'«anticipation» dans la localisation d'une limite linéaire, en fonction du stimulus tactile se déplaçant vers elle:

- 1) la réduction, telle que nous l'avons constatée au cours de ce travail, dans l'estimation de la longueur du passage d'un stimulus (sans intervention de localisation);
- 2) la réduction dans l'estimation de la durée d'un stimulus tactile, telle que nous l'avons étudiée ci-dessus;
- 3) la tendance du sujet à réduire l'amplitude quand il est sollicité de reproduire, sans le contrôle de la vue, un mouvement préalablement exécuté par lui; 2)

---

2) Les faits 3, 4, et 5 nous ont été aimablement démontrés par le Professeur A. Michotte van den Berck, de l'Université de Louvain.

- 4) la sous-estimation d'un poids donné, quand le sujet est prié de le retrouver dans une série croissante de poids;
- 5) la réaction anticipée que l'on observe quand le sujet doit établir une longueur à côté de l'étalon qu'il a sous les yeux.

Ce dernier fait a été vérifié par nous-mêmes. La longueur à reproduire était de 40 cms. Elle était présentée, à 3 mètres de distance du sujet, et à 1 m 50 du sol, sur une latte noire horizontale, de 1 m 50 de longueur, et de 10 cms. de hauteur. La longueur étalon était présentée à l'extrémité gauche sur le fond noir; elle était limitée de part et d'autre par un index vertical blanc de 8 cms de hauteur et de 12 mms de largeur. A partir de l'extrémité droite du fond noir, un troisième index se déplaçait lentement vers la gauche, sans intervention du sujet. Celui-ci avait à dire, «stop» au moment où il estimait que l'index mobile limitait la longueur désirée, à droite de l'index fixe.

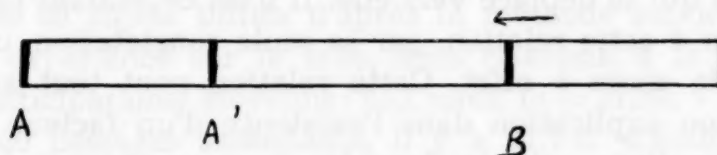


Fig. 2. A, A' = index fixes  
B = index mobile

Les expériences ont été exécutées sur cent sujets différents.

Chacun d'eux était soumis trois fois consécutivement à la même expérience. La première a donné une «anticipation» dans 76 % des cas. Moyenne: 2,6 cms ( $\sigma = 1,67$ ;  $v = 0,64$ ). La seconde nous a fourni une «anticipation» dans 67 % des cas. Moyenne: 2,1 cms ( $\sigma = 1,30$ ;  $v = 0,62$ ). La troisième a donné une «anticipation» dans 66 % des cas. Moyenne: 2,2 cms ( $\sigma = 1,42$ ;  $v = 0,65$ ). Dans la première expérience nous avons relevé 18 transgressions d'un moyenne de 1,3 cm. Dans la seconde, les 26 transgressions étaient en moyenne de 2 cms, et dans la troisième, les 29 transgressions étaient en moyenne de 1,7 cm. Des expériences homologues ont été faites de gauche à droite (la distance-étalon étant donc présentée à l'extrémité droite de la latte), chez 100 sujets dont 30 seulement avaient servi à l'expérience précédente. Au premier passage, nous avons enregistré 72 % d'«anticipations». Moyenne 2,9 cms ( $\sigma = 2,22$ ;  $v = 0,76$ ). Au second passage nous avons relevé 67 % d'«anticipations». Moyenne



ne: 2,4 cms ( $\sigma = 1,55$ ;  $v = 0,65$ ). Au dernier passage, 66 % d'«anticipations». Moyenne 2,6 cms ( $\sigma = 1,69$ ;  $v = 0,73$ ). Le nombre de transgressions, pour chacun des passages, était respectivement de 22 (Moyenne: 1,9 cms), 26 (Moyenne: 1,4 cm), et 19 (Moyenne: 1,6 cms).

Cette recherche établit une fois de plus la tendance à «anticiper» quand le sujet est sollicité de reproduire, à l'aide d'un curseur, une longueur qu'il a sous les yeux.

On peut se demander si dans ces différentes manifestations relatées, le facteur commun ne serait pas l'*attitude expectative* du sujet.

Il est vrai cependant que l'on peut mentionner une série de situations où le sujet, malgré une attitude expectative indubitable, présente une tendance à réagir d'une façon inverse. Citons entre-autres:

- 1) l'établissement du seuil extensif, en esthésiométrie, tant au cours de l'augmentation progressive de la distance entre les deux pointes, qu'au cours de la diminution progressive de cette distance;
- 2) la recherche du seuil auditif absolu, par la méthode de l'augmentation et de la diminution progressive du stimulus;
- 3) l'établissement du champ visuel, par le déplacement du «test» dans le sens centrifuge et dans le sens centripète, chez des sujets normaux ou porteurs d'un rétrécissement organique du champ visuel où, tout autant que dans les deux faits précédents, les limites indiquées par le sujet chevauchent;
- 4) l'établissement de la limite d'une anesthésie organique causée par une lésion périphérique. Nous avons en effet démontré ailleurs, que les limites indiquées par un sujet atteint d'une telle anesthésie chevauchent quand la délimitation doit se faire en raison d'un stimulus tactile qui se déplace vers cette limite, tantôt à partir de la région anesthésique, tantôt à partir de la région normalement sensible.

L'attitude expectative intervient donc dans les deux groupes de faits que nous venons de distinguer. Les résultats opposés que caractérisent chacun de ces groupes nous enseignent, que l'attitude expectative ne peut pas être la seule explication du

phénomène de l'«anticipation». Nous constatons toutefois que le conditionnement psychologique de la réaction du sujet est différente dans chacun des deux groupes de faits. Dans le premier groupe, le sujet attend le moment où il a à réagir, mais il ne trouve pas de signal sensitif concret pour sa réaction; il y a absence d'apparition ou de disparition d'une sensation ou impossibilité d'évaluer d'une façon suffisamment précise la longueur qui lui est présentée. En d'autres mots, pour pouvoir trouver le moment où il doit exécuter la consigne, le sujet tâche de faire appel à une représentation spatiale schématisée, ou bien au souvenir vague d'une limite spatiale ou temporelle ou d'une intensité pondérale, ou bien, dans le cas de la détermination d'une longueur à côté d'une autre qu'il a sous les yeux, à une perception spatiale insuffisamment précise.

Pour certaines de nos expériences il y a même des sujets qui nous ont affirmé n'avoir trouvé aucun repère et avoir été obligés de se contenter d'une «supposition» ou de «deviner».

Par contre, dans les expériences du second groupe, le sujet attend, pour réagir, le signal de l'apparition ou de la disparition d'une sensation.

Nous sommes d'avis dès lors que, dans le phénomène d'anticipation que nous avons étudié (et dans les modes de réaction semblables), l'action déterminante de l'attitude expectative peut être admise parce que conjointement le sujet se trouve dans l'impossibilité de se référer à un signal sensitif.

#### RESUMÉ

Si l'on sollicite un sujet normal de localiser une ligne tracée sur la peau, en fonction d'un stimulus tactile se déplaçant vers cette ligne, ce sujet présente une tendance marquée à le faire avant que le but ne soit atteint par le stimulus. En d'autres mots, il «anticipe» la localisation.

Ce phénomène d'«anticipation» s'est vérifié dans la très grande majorité des 720 sujets normaux adultes chez lesquels il a été exploré.

Dans le présent travail, des facteurs déterminants et des facteurs adjuvants de ce phénomène sont étudiés dans des conditions expérimentales déterminées.

Comme facteurs *déterminants* sont envisagés et étudiés: 1) la

tendance à un mode de localisation; 2) l'influence de certains repères sur le mode de localisation; 3) le mode de représentation des dimensions corporelles; 4) l'appréciation du temps de passage du stimulus nécessaire pour atteindre la ligne; 5) le mode d'estimation de la longueur du déplacement du stimulus tactile.

Comme facteurs *adjuvants* sont explorés: 1) la nature du stimulus; 2) la vitesse du déplacement du stimulus; 3) la longueur du trajet de ce stimulus; 4) les degrés de schématisation de la ligne à localiser.

L'étude expérimentale des différents facteurs considérés comme adjuvants fournit des données positives quant à leur action. Par contre, des facteurs déterminants envisagés, seul le mode d'estimation de la longueur du trajet du stimulus tactile se montre en relation avec le phénomène de l'«anticipation» dans la localisation d'une limite linéaire.

Il n'est cependant pas permis d'attribuer à cette relation, par sa seule constatation, une signification de cause à effet. Cette relation peut tout aussi bien signifier l'existence d'un facteur commun à l'«anticipation» relevée dans les deux expériences.

Dans le même ordre d'idées, vient à l'esprit la question de savoir si le phénomène de l'«anticipation» ne se rattache pas à un ensemble de manifestations semblables qui auraient un facteur déterminant commun et si, dans ces différentes manifestations, ce facteur commun ne serait pas l'*attitude expectative*.

Malheureusement, l'on peut mentionner une série de situations où le sujet, malgré une attitude expectative indubitable, présente une tendance à réagir d'une façon inverse de l'«anticipation». Les résultats opposés qui caractérisent chacun de ces groupes de faits nous enseignent que l'attitude expectative ne peut pas être la seule explication du phénomène de l'«anticipation».

Nous constatons toutefois que, dans le premier groupe de faits, le sujet attend le moment où il a à réagir, mais qu'il ne dispose pas de signal sensitif concret pour sa réaction.

Par contre, dans les expériences du second groupe, le sujet attend et trouve pour réagir le signal de l'apparition ou de la disparition d'une sensation.

Nous sommes d'avis dès lors, que dans le phénomène d'«anticipation» étudié dans le présent travail (et dans les modes de réaction semblables) l'action déterminante de l'attitude expectative

tative peut être admise, parce que conjointement le sujet se trouve dans l'impossibilité de se référer à un signal sensitif.

#### SUMMARY

When a normal subject is asked to localize a line drawn in the skin, in function of a tactile stimulus which is being moved towards this line, this subject shows a pronounced tendency to do so before the aim is reached by the stimulus. With other words, he "anticipates" the localization.

This phenomenon of "anticipation" has proved correct in the very great majority of the 720 normal adult subjects, upon whom it has been examined.

In the present work, the *determining factors* and the *auxiliary ones* of this phenomenon have been studied under well defined experimental conditions.

As *determining factors* have been taken into consideration and are studied: 1) the tendency towards a mode of localization; 2) the influence of certain guide-marks upon the mode of localizing; 3) the mode of representation of bodily dimensions; 4) the appreciation of the time of the moving of the stimulus necessary to reach the line; 5) the mode of estimating the length of the movement of the tactile stimulus.

As *auxiliary factors* have been examined: 1) the nature of the stimulus; 2) the moving speed of the stimulus; 3) the length of the course of the stimulus; 4) the degrees of schematization of the line to be localized.

The experimental study of the different factors which are considered as auxiliary ones, supplies us with positive data concerning their action. On the other hand however, among the determining factors which were taken into consideration, only the mode of estimating the length of the movement of the tactile stimulus shows a relation with the phenomenon of "anticipation" in the act of localizing a linear limit.

It is not allowed however to assert that this relation, only because of its statements, should imply that one phenomenon is the cause of the other. This relation may as well mean that both cases of "anticipation" have a common cause.

This leads us to ask whether the phenomenon of "anticipation" does not belong to a group of similar phenomena, that



should have a common determining factor, and, if so, whether this common factor is not the *expectant attitude*.

Unfortunately, it is possible to quote a series of situations in which the subject, notwithstanding an indubitable expectant attitude, shows a tendency to react opposite to the "anticipation".

The opposite results met with in each of these groups of facts teach us that the *expectant attitude* may not be the only explanation of the phenomenon of "anticipation".

We however observe that, in one of both groups of facts, the subject awaits the moment upon which is to react, but that he does not meet a sensitive sign for his reaction.

On the other hand, in the experiments of the other group, the subject awaits and finds, in order to react, the appearing or disappearing of a sensitive sign.

We therefore are of opinion that, in the phenomenon of "anticipation" which has been studied in the present work (and in the similar modes of reaction), the determining action of the *expectant attitude* may be admitted, because, at the same time, the subject is given no possibility of referring to a sensitive sign.

#### BIBLIOGRAPHIE

- Hall, G. S. and H. H. Donaldson, Motor Sensation on the skin. *Mind*. X. 1885.
- Henri, V., Recherches sur la localisation des sensations tactiles. — *Année psychol.* 2, 1895.
- Kries, J. von, *Allgemeine Sinnesphysiologie*. — Leipzig, 1923.
- Linke, F., *Grundfragen der Wahrnehmungslehre*. — Munich, 1929.
- Montgomery, E., *Space and Touch-Mind*, X, 1885.
- Nyssen R. et J. Hozay, La délimitation subjective des anesthésies cutanées tactiles — *Folia, Psychiatr., Neurol. et Neuroch. Neerlandica*, 53, 4 et 5, 1950.
- , De la délimitation des régions cutanées par la méthode du passage d'un stimulus tactile. *Acta Psychologica*. — VIII, 2, 1951.
- Pillsbury, W., Some questions of cutaneous sensibility. *Amer. J. of Psychol.*, 1895.
- Spearman, C., Einfluss der Bewegungsrichtung auf den Lokalisationsfehler. *Phil. Stud.*, 1906.
- Wundt, W., *Grundzüge der Physiologischen Psychologie*. Vol. II, 1902, Leipzig.
- Zigler, M., The experimental relation of the two-point limen to the error of localization. — *J. gen. Psychol.* 1935.

## SOCIAL THINKING

BY

JOHN COHEN

### (i) INTRODUCTION

We can take as our starting point a remark made by Schrödinger in his essay *Science and Humanism*. (17) 'Who', he asks, 'would expect originality from a committee or commission or board or that sort of thing?' <sup>1)</sup> In the last few decades there has been an immense growth in the number of bodies whose task requires collective or organized group thinking. This is true of industry, the professions, and of municipal, national and international affairs. In Gt. Britain alone the number of committees must be of the order of several hundred thousand. Thousands of sittings of organs of the United Nations are held every year. Are we to assume that all this activity is totally devoid of originality and that, in the nature of things, groups are incapable of novel thinking?

Groups and institutions vary in many ways which may have a bearing on their capacity for original achievement. One way is the manner in which they are first set up. Some committees, for example, generate themselves. Others are appointed, and the selective factor which determines the choice of members is decisive in fixing the outcome of the deliberations. Institutions vary in the strictness of their administrating control. In some, innovation may be encouraged and information freely disseminated. In others ideas are transmitted down the hierarchy of authority, and anyone who suggests a new way of doing things is looked at with suspicion.

Unfortunately the available evidence does not allow of a definite answer to Schrödinger's rhetorical question. I shall therefore confine myself in what follows to a consideration

---

<sup>1)</sup> This view is not held by Schrödinger alone. 'It is probable that the most significant innovations will come from one rather than from a team'. Brosin, H. W. in Jeffress L. A. (ed.) *Cerebral Mechanisms in Behaviour*, New York: Wiley 1951, p. 292.

of the results of recent research in to the various factors that influence social thinking, and clues to an answer to Schrödinger's question may arise therefrom. By 'social thinking' I shall mean the intellectual activity of committees, discussion groups, conferences, and social institutions. I shall be concerned in the main with committees and conferences.

## (ii) PREVIOUS STUDIES

Let us first glance very briefly at the results of early studies in this field. These were chiefly attempts to examine the effects of group discussion on the previous judgements of individuals. In the typical experiment separate individuals are first given a task the accuracy of which can be independently estimated. They are then formed into groups to discuss the judgments previously made by the constituent members separately. The comparative accuracy of the individual and group judgments is then assessed.

One of the earliest investigators seems to have been Münsterberg (15) in 1914. His subjects were given cards carrying a large number of dots, some cards having more dots than others. Each subject had to select the card which seemed to have the most dots. After the individual judgments were made, the subjects were formed into groups who discussed the accuracy of the individual estimates. His results showed that group discussion has the effect of bringing about much more accurate judgments than those reached by individuals. There is a tendency in group situations, as F. H. Allport (1) found in 1920, to avoid extremes even though the judgments are not made aloud; and individuals tend to converge towards a norm. In 1924 Bekhterev and Lange (3) carried out similar experiments and arrived at much the same conclusions. A number of objects were exposed for 15 seconds and subjects had to record what they observed. Group discussion then took place. As a result of discussion the proportion of correct observations increased, more correct than incorrect items being added. South, (19) in 1927, found that small groups are superior to large ones when the material lends itself to immediate formation of opinion, and larger ones are better when it is desirable quickly to reject wrong proposals or solutions. A year later Watson (22) showed

that the superiority of groups over individuals varies with the kind of task. Group superiority is more marked when high performance depends on the summation of many responses or on much inter-stimulation. Shaw's (13) experiment in 1932 had a rather more elaborate design. Five groups of four persons worked on a number of problems in separate rooms. A control series of twenty-one subjects worked on the same problems individually. Two weeks later the experiment was repeated with a different set of problems, only this time those who had worked in groups worked individually and *vice versa*. Some 53 per cent correct solutions were obtained by the groups as compared with about 8 per cent obtained by the individual subjects. Shaw explains this difference on the grounds that in a group situation there are more ways of considering a problem, more solutions are offered and each receives more criticism than when individuals work in isolation, and there is more likelihood of yielding to the criticism of another person than to one's own. Jeness (12) in 1932 introduced a change in the procedure of this sort of experiment. His subjects had to estimate the number of beans in a jar. He then divided the subjects into two groups, one composed of individuals who had given very similar estimates and the other composed of individuals who had given very different estimates. In the first group the average error of the individual estimates declined only by about 17 per cent as compared with a decline of about 60 per cent in the second group.

The foregoing experiments are consistent in showing that, in the performance of certain limited tasks, groups are superior to individuals, the superiority revealing itself in greater accuracy or in greater success in problem-solving. They were designed to answer rather specific questions and were not related to any body of theory or intended to assess the usefulness of a particular theoretical model. As experiments, they are not specially relevant to the usual functions of groups. A conference is not normally convened *simply* because its conclusions are likely to be more accurate than would be the 'average conclusion' of the participants were they to meditate in isolation of one another. Nor is it usually convened to find the solution to clear-cut problems. It is primarily designed to provide the opportunity for inter-communication between the participants to an end which is *external* of the conference process itself. The superior



accuracy or skill in problem-solving (in the narrow sense) of group over individual thinking is not by itself an adequate test of effectiveness.

### (iii) CRITERIA OF EFFECTIVENESS

How then may the effectiveness of group activity be assessed? Two kinds of criterion may be distinguished, internal and external. An internal criterion in a social system is one based entirely on effects within the system. An external criterion is one which is based on control of the system by some agency which acts outside it (8). The use of internal criteria alone for controlling the course of events in a social system leads to circular action; there is no link with the outside. A conference has to know in which direction it is 'moving' and with what speed. Reliance on internal criteria simply assures that some activity has taken place. A feeling of euphoria after a meeting is like the state of mind of the genial captain 'who hears that his course is too much to the left, rushes to the wheel, turns it to the right and, having done so, goes happily to dinner. In the meantime his boat goes round in circles'. (13)

The importance of this distinction between internal and external criteria does not seem to have been appreciated in recent research in this field. J. G. Darley (10), for example, suggests that a 'constant search for an external criterion' may be setting too difficult a goal, and possible 'all we need or can get in our early studies of group behaviour is a criterion that adequately subsumes the satisfactions of the individuals in the group'. This seems to be the crux of the matter. It is precisely on the correct choice and use of external criteria that the measurement of social thinking depends.

The limitations of the use of internal criteria may be illustrated by the investigations recently reported by Marquis (14) and his associates. In this excellent study 72 conference groups in business, industry and government were subjected to close examination. The size of the groups ranged from 5 to 17 and all of them had the task of making decisions. The criteria of effectiveness employed were (i) members' satisfaction with the meeting, (ii) productivity of the meetings as judged by the

number of items on the agenda completed, and (iii) extent of disagreement with the decision at the end of the meetings. The data collected by the observers related in the main to the satisfaction and amount and kind of participation of members of the groups. A composite 'index of cohesiveness' was devised which could be regarded as measuring the attractiveness of the group for its members. A correlation was found between this index and members' satisfaction with the meeting, but *not* with productivity.

The use of external criteria presupposes that conferences, far from being closed systems, are brought into being so as to produce a change in the wider world of which they are part. One external test of the effectiveness of a conference of management and workers, for example might be the degree of improvement in subsequent industrial relations. Costing and accounting procedures are forms of external economic test. Duration of stay of patients in hospital (6) and relapse rates are external tests of the effectiveness of institutions for medical treatment. Rates of recidivism constitute one test of the effectiveness of detention in prisons. In general, for social institutions to be effective there must be a link between planner and executive and a feedback from the second to the first (13).

External criteria in the form of inverse feedback, however, do no more than maintain a stable system. If we look for that originality in social thinking which Schrödinger misses we shall need methods of *positive* feedback. Grey Walter (23) has already suggested that 'limited positive feedback is the system of choice for exploration and enterprise'. The task is accordingly to develop such conditions in group situations. Obstacles to communication and receptivity would presumably have to be minimised. Entirely novel ideas are likely to arise when participants are optimally receptive and responsive to continuously changing patterns of events which are functions of the group setting. One possibility in this connection presents itself from a different area of psychological enquiry, namely, experiments on the 'level of aspiration'. This is defined as the level of performance which a person sets himself in a given task when he knows his past level of performance in that task. Experiments show that in given circumstances success tends to raise the level of aspiration. In other words, the effect of success on aspiration has the character

of a positive feedback. It would seem to follow that in group situations where it is desired that every participant should exert his best efforts, the tasks assigned to each should be close to each individual's true level of aspiration.

#### (iv) OBSTACLES TO COMMUNICATION

We have now to consider some of the more important factors which can be identified as hampering the process of communication among members of a group (8). We should probably include: undue homogeneity of the members; the imposition of demands by a higher authority which the group members are not fitted or emotionally prepared to meet; the tradition of resisting any *new* proposal even when no good reason can be given for rejecting it; and stereotyped notions of the supposed roles of members.<sup>2)</sup> There are no doubt many other perhaps more obstructive influences, but the factor which perhaps accounts for most of the variance of inhibition in communicating appears to be the tacit assumption implied in the behaviour of the participants that agreement is desirable as an end in itself. This is an internal criterion of group effectiveness. It tends to lead the group away from the exploration of differences, which is possibly the primary task, to a discussion of areas of agreement. The total pattern of group behaviour might be interpreted as implying a tendency to conform to the group and yield to its pressures, a *compulsion to agree*. To a greater or less

---

<sup>2)</sup> The following example illustrates how this can happen. A committee which had completed its work and was about to disperse received a request from a higher authority to submit further recommendations on a certain subject. This request came as a complete surprise to all the members except one who had known all along of the committee's obligations to submit recommendations to the authority in question. Had the awareness of this obligation been general, members would certainly have devoted time and thought to devising ways of fulfilling it. At the eleventh hour it was too late to prepare any useful material. When the one who had known was asked: "Why did you not tell us before?", the reply came, perhaps half in jest, "No one asked me". The belief that it is possible to defend negligence with the plea that one had not been asked had serious consequences. If the social climate had been such that the excuse 'No one asked me' would be unthinkable and unacceptable in any circumstances, a situation like this could never have arisen.

extent, varying with the circumstances and the individuals, attention is turned from points on which members diverge. There is a retraction of the social field to exclude whatever is nonconforming or provokes anxiety. The chief effect is thus to discourage a free and uninhibited expression of view. It is rather surprising to find the formula: 'Discussion is splendid but disagreement is not' sometimes receives explicit approval from psychologists (21).<sup>3)</sup> The nearer a group moves towards a decision, the more any divergence threatens its solidarity. The compulsion to agree therefore seems to grow stronger as the time for decision approaches, a hypothesis which could be put to experimental test. Preoccupation with agreement precludes an effort to listen and understand what another person is saying. One agrees or disagrees with what one does not understand. I do not mean that no differences of view are expressed or that stormy meetings are unknown, but that, generally speaking, differences are kept strictly under control. Even when an attempt is made by a member to express a divergent point of view the effort is usually abandoned in the end. One or other party, often the less voluble or less dominant, surrenders. The 'elasticity' of the group, in the sense of its tolerance of differences without disintegrating, is accordingly limited. The compulsion to *disagree* also occurs, but it apparently seems to be much more rare.

I have elsewhere (9) suggested that the explanation for this substitution of agreement for understanding is perhaps ultimately to be sought in our child-rearing and educational practices. The child rarely has the experience of being understood rather than being controlled by someone else. Trying to understand another may be regarded as the counterpart of being understood. Both these elements seem to be lacking in our culture pattern. Among the techniques that might be applied for improving mutual understanding, the method of reversing or exchanging roles in a dispute perhaps deserves careful study (7).

---

<sup>3)</sup> A remark of Conant is of interest in this connection: 'It seems that in the embryonic stages of each of the modern disciplines, violent polemics rather than reasoned opinions often flowed most easily from the pen' J. B. Conant, *On Understanding Science*, London: Oxford University Press, 1947, p. 6.



Judged by external criteria of effectiveness there is not necessarily any merit or advantage in agreement as such. If it is a substitute for the exploration of differences it clearly reduces the possibility of interaction among members. The interaction on which group thinking depends cannot take place when all participants tend to think alike. Often agreement may be simply 'a device for maintaining the illusion of one's omniscience'. In scientific discussion, as W. B. Cannon (5) has remarked, disputation may do more good than harm. He points out that the discovery of sympathin and the operation of complete sympathectomy were the direct consequences of a polemic. In the work of public bodies like Royal Commissions, as illustrated by the Webbs' famous minority report on the Poor Law, divergence from the group may in the long run prove far the more fruitful course to take.

The hypothesis of a compulsion to agree is supported by recent experiments of S. E. Asch (2) at Swathmore College. His method is to place an individual in a relation of conflict with most or all the members of a group and then record the effects on the individual's judgment. For example, the group members are asked to match the length of a given line with one of three unequal lines, one of which seems to the naive observer to be similar in length to the given line. One member finds himself contradicted by the rest, who are instructed beforehand to give false judgments. Each member announces his judgment publicly. The effect of the majority in compelling agreement on the part of the 'critical' individual who finds himself in the minority is measured by the frequency of errors in the direction of the distorted estimates of the majority. The results of these experiments show that about a third of the 'critical' individuals found themselves compelled to agree with the majority. This effect becomes more marked as the situation is less clear. On being interviewed afterwards these individuals confessed to have been disoriented, doubt-ridden or 'experienced a powerful impulse not to appear different from the majority'.

A number of experimental variations have been introduced. For example, in order to study the effect of a non-unanimous majority, one of the instructed group is asked to deviate in prescribed ways. The result is to reduce appreciably the majority effect. A unanimous majority of three is more compelling than

a majority of seven and one dissenter. The compulsion to agree seems to diminish when there is some support for divergence.

The fact that in the study of Marquis referred to above there was no correlation between his second and third criteria i.e. between productivity and amount of residual agreement, suggests that agreement *per se* is not necessary for productivity, as defined by him. This would seem to argue in favour of the need for understanding differences rather than for mere agreement without any assurance that differences have been resolved.

Further light is thrown on this problem by experiments of Festinger (11) and his fellow-workers at the University of Michigan. These experiments are concerned with the analysis of factors determining communication in groups. One hypothesis, for example, supposes that the frequency of communication between members of a group increases with greater discrepancy of viewpoint. This has been tested by 'planting' a member in a group with instructions to maintain a divergent opinion. In one experiment this member received five times as many communications as the other participants. This suggests that the disclosure of differences, at any rate up to a certain point, stimulates interaction. When discrepancy of viewpoint is held constant, communication on a given item appears to vary with the perceived relevance of the item. This hypothesis has been tested in experiments in which an item is arbitrarily made relevant for some members and irrelevant for others. A number of other hypotheses relating to various alleged sources of pressure to communicate have also been tentatively examined.

#### (v) POSSIBLE LIMITATIONS OF A CYBERNETIC MODEL

It has seemed quite natural in the preceding discussion to regard a group engaged in social thinking as a system of communication. The use of a general analogy or model from cybernetics is at least helpful. But there are dangers of oversimplification in such a procedure. Russell Brain's (4) caveat in relation to such analogies with the individual mind has even more force in relation to group phenomena. 'When the electrical experts', writes Dr. Brain, 'claim to have detected in the physiology of the nervous system a likeness to their machines,

they may or may not be right, but if they try to express that likeness in terms of the mind, they are guessing the meaning of a language of which they have not yet learned the alphabet' <sup>4)</sup>. The psychological complexities are considerably greater in a group than in any individual situation and the limitations of analogies from other realms must be all the more appreciated.

An investigator from Mars would know very little about our social relationships if he merely studied the channels of communication in the postal and telephone services without opening letters and tapping messages. Apart from the complexity of the channels themselves, there is also the content communicated. A person may send an ambivalent message consciously intending one meaning and unconsciously another. The receiver may likewise receive the message ambivalently. One and the same message, such as a remark addressed to several people at a meeting, may be received with different meanings by the several listeners. The manifest content of a message must be distinguished from its latent content as illustrated, for example, in forms of irony, in which instead of saying what one thinks, one pretends to think what one says.

In any group, communication may be vocal, not simply verbal (20). In vocal as distinct from verbal communication there are such factors as intonation, emphasis, rate of speech, and ease or difficulty of articulation. It is perhaps safe to assume that the more conventionally a statement is expressed the less confidence we are entitled to have in our interpretation of it. There may be non-vocal communication as well. A person's silence may have one or more of several possible meanings. He may wish it to convey one meaning to some and another meaning to others present. Or he may wish to conceal the true meaning of his silence and camouflage it subsequently when he eventually speaks.

Communication may be affected by so-called 'parataxic' distortion. One's perception of person may be so distorted for any length of time. When this happens, the other person's 'real' characteristics, as defined, say, by a consensus of competent observers who know him well, are unconsciously modified by attributing to him imaginary characteristics, by seeing him not

---

<sup>4)</sup> See also G. Jefferson, *Brit. Med. J.*, 1949, 1, 1105.

as he is but as endowed with qualities which the perceiver has previously associated with certain other persons (20). In view of these considerations it would seem that designers of a model for a brain have some way to go before they reach the point when they can produce a model capable of simulating the complexities of group thinking.

A cybernetic model, though it may eventually prove very useful, is not the only one that may be used in the study of group phenomena. Penrose (16) has recently shown that fruitful methods of experiment and statistical analysis may be developed in the study of groups by taking a model from epidemiology. On the assumption that an individual possessing a given 'idea' or 'reaction pattern' may be taken as the fundamental unit, he has shown, for example, that it is possible with the help of Gaussian tables to measure the power exerted by a resolute minority over a majority composed of individuals behaving more or less at random, as may happen in voting situations. He suggests, furthermore, that in the transmission of ideas in 'mental epidemiology', the three factors (i) quality of the idea, (ii) means of transmission, and (iii) state of the recipient, may be regarded as corresponding to the three commonly accepted factors in physical epidemiology, viz. (i) the infective agent, (ii) means of transmission, and (iii) susceptibility of the exposed population. On this basis he proceeds to analyse the temporal course followed in the transmission of ideas in crazes, panic, religious outbreaks, and politics. These concepts and methods are likely to prove most valuable in stimulating further research in this elusive field of study.

#### (vi) CONCLUSION

Let us now return to our original question. It would seem that too little is yet known of the nature of interaction in groups to justify our dismissing the possibility of their having a capacity for original thinking. In principle, the thinking of an individual is always bound up with 'invisible' groups. His thoughts are not generated spontaneously — they invariably have a social as well as a personal history. The actual presence of a group, provided there is a complete freedom from inhibition and optimal receptivity, may be regarded as enabling us to see a



history of ideas in the making. It is true, of course, that the actual emergence of a new idea must take place in an individual mind. But when a person participates in a group there are possibilities of assimilating the thoughts of others and of accommodating to them which cannot exist when he works alone. In short, there are no *a priori* grounds for despising the group as a vehicle of original thought. Only further experiment can lead to a true evaluation of group resources for intellectual innovation.

### SUMMARY

An attempt is made to determine, in the light of the evidence, whether original thinking can be achieved by groups such as committees and conferences. The work of early investigators suggests that groups are superior to individuals in the performance of certain limited tasks. More elaborate recent studies are criticised for failing (i) to distinguish clearly between internal and external criteria of group effectiveness and (ii) to appreciate the role of external criteria in assessing such effectiveness.

Various obstacles to inter-group communication are discussed, in particular, the 'compulsion to agree', the source of which, it is suggested, may be traced to educational and child-rearing practices in the culture. The limitations of a cybernetic model and the uses of an epidemiological model for certain problems are briefly indicated. The upshot is that we cannot rule out the possibility of originality in groups. On the contrary, provided members are receptive, group conditions may be particularly favourable to innovation. This hypothesis requires to be tested experimentally.

### REFERENCES

1. Allport, F. H., The influence of the group in association and thought, *J. Exper. Psychol.*, 1920, 3: 159—182.
2. Asch, S. E., Effects of group pressure upon the modification and distortion of judgments, pp. 170—190 in Guetzkow, H. (ed.) *Groups, Leadership and Men*, Pittsburgh, Pa., Carnegie Press, 1951.
3. Bekhterev, W. and M. Lange, Die Ergebnisse des Experiments auf dem Gebiete der Kollektiven Reflexologie, *Zeitschr. für die angew. Psychol.*, 1924, 24: 224—254.

4. Brain, W. Russell, *Mind, Perception and Science*, Oxford: Blackwell, 1951, p. 86.
5. Cannon, W. B., *The Way of an Investigator*, New York: Norton, 1945, pp. 105—6.
6. Cohen, John., *The Recruitment and Training of Nurses*, London: His Majesty's Stationery Office, 1948.
7. ———, The technique of role reversal, *Occupat. Psychol.*, 1951, 25: 64—6.
8. ———, Some working hypotheses and provisional conclusions in the study of committees and conferences, *Occupat. Psychol.*, 1952, 26: 70—7.
9. ———, The social psychology of childhood, *Biol. and Human Affairs*, 1952, 17: 183—8.
10. Darley, J. G. in Guetzkow, H. (ed.) *op cit.* p. 257.
11. Festinger, L. *et al.*, *Theory and Experiment in Social Communication*, University of Michigan, 1950.
12. Jeness, A., Social influences on change of opinion, *J. Abn. Soc. Psychol.*, 1932, 27: 29—34, 279—96.
13. Lewin, K., Frontiers in group dynamics, *Human Relations*, 1947, 1: 143—53.
14. Marquis, D. G., H. Guetzkow and R. W. Heyns, pp. 55—67 in Guetzkow, H. (ed.) *op. cit.*
15. Münsterberg, H., *Psychology of Social Sanity*, New York: Doubleday Page, 1914.
16. Penrose, L. S., *The Objective Study of Crowd Behaviour*, London: H. K. Lewis, 1952.
17. Schrödinger, E., *Science and Humanism*, London: Cambridge University Press, 1951, p. 8.
18. Shaw, M. E., A comparison of individuals and small groups in the rational solution of complex problems, *Amer. J. Psychol.*, 1932, 44: 491—504.
19. South, E. B., Some psychological aspects of committee work, *J. Appl. Psychol.*, 1927, 11: 348—68, 437—64.
20. Sullivan, H. S., The psychiatric interview, *Psychiatry*, 1952, 15.
21. Volkmann, J., Scales of judgment, in Rohrer, J. H. and Sherif, M. (eds.) *Social Psychology at the Crossroads*, New York: Harper, 1951, p. 282.
22. Watson, G. B., Do groups think more efficiently than individuals? *J. Abn. Soc. Psychol.*, 1930, 21: 79—109, (See D. Katz, The influence of the group upon social behaviour and attitudes, pp. 238—58 in Guilford, J. P. (ed.) *Fields of Psychology*, New York: Van Nostrand, 1950, sec. edit.).
23. Walter, W. Grey, Features in the electro-physiology of mental mechanisms, pp. 67—78 in Richter, D. (ed.) *Perspectives in Neuropsychiatry*, London: H. K. Lewis, 1950.

*Department of Ophthalmology, Wilhelmina Hospital,  
University of Amsterdam*

## SOME ASPECTS OF APPARENT MOTION

BY

W. P. C. ZEEMAN, M. D. and C. OTTO ROELOFS, M.D.

When a white square is projected at a place A on a wall or screen and then withdrawn, while a second, exactly similar square is projected at a place B, the observer will, under suitable conditions, not perceive any appearing and disappearing but will have the impression of a white square moving apparently without interruption from A to B. This frequently described apparent motion is known as  $\varphi$  motion.

When we are concerned with an optimal apparent motion of this kind, obtained by the use of 2 white squares, it is not possible to ascertain whether the first square disappears in the direction of B or the second appears from the direction of A; one sees only one square that moves over the whole distance A—B. If, however, the first square is projected with red light and the second with green light, the observer sees the red square disappearing in the direction B and the green one appearing from the direction A.

This movement of both the red and the green square indicates a labile localization both in the appearance and in the disappearance of luminous stimuli. This labile localization makes identification of both squares possible. It is a known fact that the egocentric localization (often labelled with the incorrect name of absolute localization) is on the whole highly labile and uncertain, whereas the exocentric localization (also frequently termed relative localization) is very stable.

It appears not impossible that with the successive projection of 2 squares the observer depends chiefly on the labile egocentric localization, which might facilitate identification of the 2 squares.

One of the conditions that must be fulfilled for optimal apparent motion is that places A and B must not lie too far

apart. The duration of the interval between disappearance of the first object and appearance of the second is also very important. In connection with the foregoing, it may be expected that a stabilization of localization of each of the objects will oppose an optimal apparent motion. The investigation described below constitutes an attempt to establish this presumption. The investigation also brought to light certain aspects of apparent motion which were unknown to us and which appear worthy of further consideration.

The experiments were carried out as follows: The test subject (observer) was seated at a distance of 2.5 m from the screen. The projected white square had an area of  $10 \times 10$  cm. Each side of the square thus subtended an angle of  $2^\circ 13' 30''$  at the eye. The distance from A to B was varied: 80, 60, 40, 20 or 10 cm. The projection at B followed that at A without a pause. The square was not simply projected once at A and once at B but with continual alternation from A to B and from B back to A. The following durations of each individual projection were employed:  $\frac{1}{6}$  sec. (167 msec.),  $\frac{1}{4}$  sec. (250 msec.),  $\frac{1}{3}$  sec. (333 msec.),  $\frac{1}{2}$  sec. (500 msec.),  $\frac{2}{3}$  sec. (667 msec.), 1 sec. (1,000 msec.),  $\frac{5}{4}$  sec. (1,250 msec.),  $\frac{5}{3}$  sec. (1,667 msec.) and  $\frac{5}{2}$  sec. (2,500 msec.).

We shall now discuss in turn the observations made by us under these different conditions, comparing first the results obtained when the distance A—B was kept constant and only the times of projection were varied.

GROUP I. The distance from A to B was 80 cm or  $18^\circ 10' 50''$  (see graph. I).

*Projection time  $\frac{1}{6}$  sec.* The observer sees 2 squares flickering with rapidly varying light intensity. There is no question of any apparent motion of the squares. Upon fixating one of the squares one often gets the impression that the rate of fluctuation of the light intensity of the non-fixated square is slower than that of the fixated square. If one looks at the black space between the places of projection of the 2 squares, one does perceive motion here. When B appears one sees a displacement of black in the direction of A, and when A appears one sees a displacement of black towards B. Here we may ask whether it is really the intervening black that appears to move or the whole background against which the luminous phenomena take place. There is no



doubt that the former is the case. This displacement of black takes place in the region *between* the 2 intermittent light

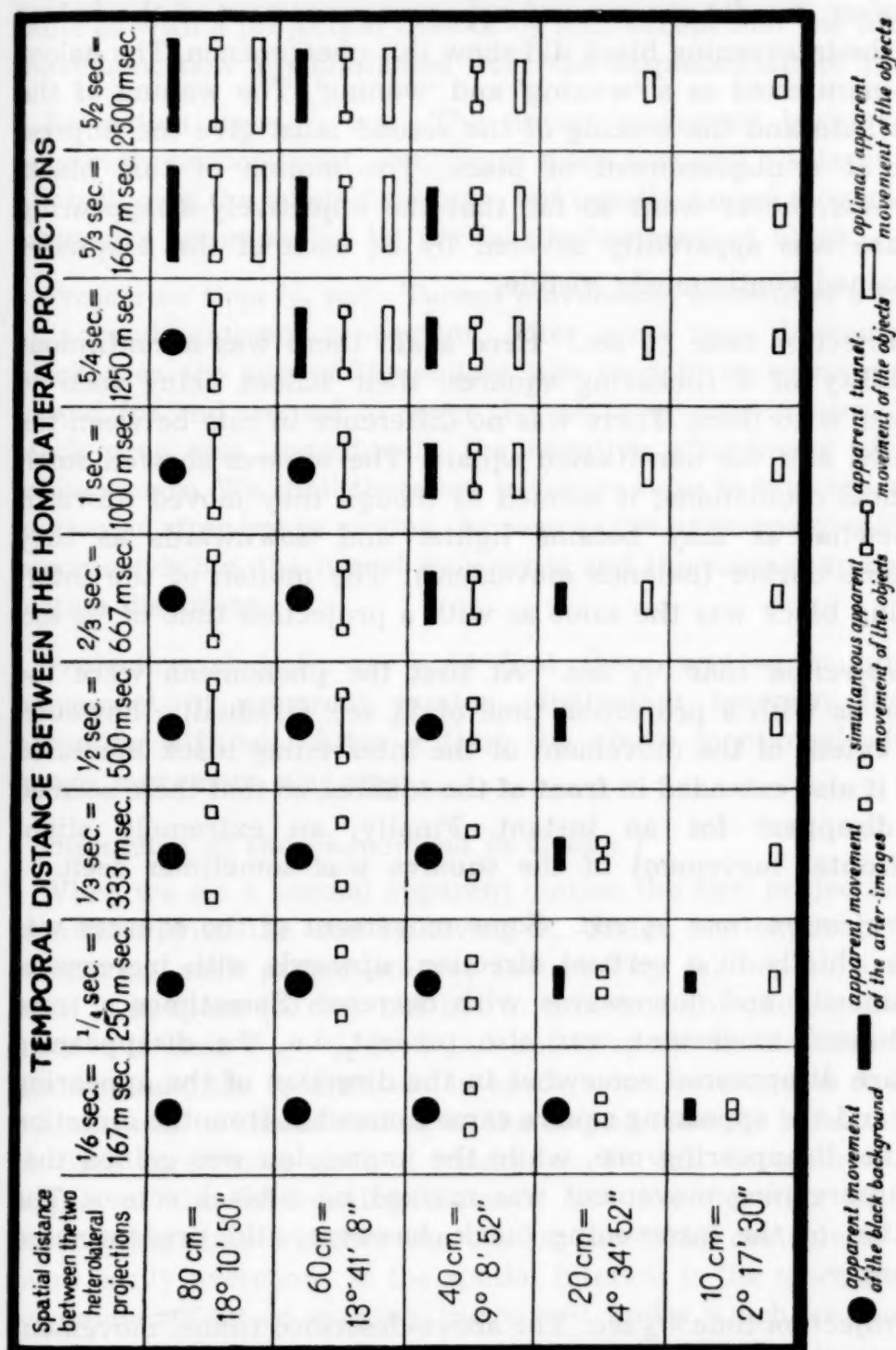


Fig. 1.

phenomena and is probably connected with the very distinct halo seen around each square. The haloes come much closer to

each other than the squares themselves, so that an apparent motion of the former might be expected. Strange to say, however, we did not see any apparent movement of the haloes but the intervening black did show this phenomenon. The haloes are manifested as a 'waxing' and 'waning'. The waning of the first halo and the waxing of the second must give the impression of a displacement of black. The motion of this black, however, never went so far that the objectively-disappearing square was apparently covered by it; each of the 2 squares remained continuously visible.

*Projection time  $\frac{1}{4}$  sec.* Here again there was a continuous visibility of 2 flickering squares, their haloes being seen to flicker with them. There was no difference in rate between the fixated and the non-fixated square. The squares showed small vertical oscillations; it seemed as though they moved upwards somewhat as they became lighter and downwards as they became darker (balance movement). The motion of the intervening black was the same as with a projection time of  $\frac{1}{6}$  sec.

*Projection time  $\frac{1}{3}$  sec.* At first the phenomena were the same as with a projection time of  $\frac{1}{4}$  sec. Gradually, however, the extent of the movement of the intervening black increased and it also extended in front of the squares, so that these seemed to disappear for an instant. Finally, an extremely slight horizontal movement of the squares was sometimes seen.

*Projection time  $\frac{1}{2}$  sec.* Some movement of the squares was seen, chiefly in a vertical direction: upwards with increase of luminosity and downwards with decrease. Sometimes a trace of tunnel movement was also present, i.e. the disappearing square disappeared somewhat in the direction of the appearing one and the appearing square came somewhat from the direction of the disappearing one, while the impression was gained that an intervening movement was masked by a black screen. The motion of the intervening black, however, still predominated strongly.

*Projection time  $\frac{2}{3}$  sec.* The above-described tunnel movement was now much more marked. There was an alternation of tunnel movement and movement of the intervening black as described for projection times of  $\frac{1}{3}$  and  $\frac{1}{2}$  sec. The displacement

of black still predominated somewhat over the tunnel movement.

*Projection time 1 sec.* The phenomena were practically the same as with a projection time of  $\frac{2}{3}$  sec., except that the tunnel movement now predominated over the displacement of black.

*Projection time  $\frac{5}{4}$  sec.* The tunnel movement became increasingly predominant; sometimes there was still displacement of black over the whole field only, but usually tunnel movement, often still accompanied by partial displacement of black.

*Projection time  $\frac{5}{3}$  sec.* Tunnel movement; sometimes a trace of normal apparent movement. After some time, however, a black area the size of the square was seen to move regularly to and fro and seemed to cover the white squares in turn. The black area was undoubtedly the negative after-image of the white square. We shall therefore in future refer to this apparent motion as after-image motion. As soon as the after-image motion becomes visible, the tunnel movement and the normal apparent motion disappear.

*Projection time  $\frac{5}{2}$  sec.* At first there was again tunnel movement of apparent motion, distinction between the 2 becoming difficult. After a time the above mentioned after-image movement was seen.

#### *Discussion of the phenomena in Group I*

When we see a normal apparent motion the first projection is identified with the one following it. Under our experimental conditions, each projection was followed by (a) one separated from it in space (80 cm) but not in time and (b) one separated from it in time ( $\frac{1}{6}$ — $\frac{5}{2}$  sec.) but not in space. In the first case the distance, the localization in space, constitutes more or less a hindrance to identification; in the second case the localization in time constitutes such a hindrance. It will depend on the length of the intervening distance and the duration of the intervening time whether the one or the other obstacle is the more easily overcome. If the spatial interval is the more easily overcome we get an apparent movement, under which we should perhaps also include the tunnel movement; if the temporal interval is the more easily overcome, the apparent motion is not seen and only a black displacement or after-image movement

remains possible. As long as the time interval between 2 projections at the same place is smaller than 333 msec. there is not even a trace of apparent motion. This is possibly connected with the duration and the decrease of intensity of the *positive after-image*. If the intensity of the positive after-image has not decreased to any considerable degree during the pause between 2 projections at same place, the light intensity will fluctuate but the square will not disappear completely.

What does the literature tell us about the duration of the positive after-image? According to Helmholtz the positive after-image attains its greatest brightness after an illumination time of  $\frac{1}{3}$  sec. The greater the intensity of the primary light, the greater will be the intensity of the positive after-image and the longer its duration. Even with weakly illuminated objects it is still possible to get quite good positive after-images with a measurable duration. The duration is believed to range from 2 sec. to several minutes (sunlight). We are not interested so much in the total duration of the positive after-image as in the time during which its intensity can still exert an influence. The flicker frequency is sufficient proof that the intensity decreases appreciably after only a very short time. The question is, however, after how long the intensity has decreased so far that — as in our experiments — it must have lost its significance for the identification. This will be at the moment that the site of the projection looks practically black again after disappearance of the light. An investigation carried out by Plateau seems to give hints in this direction. A disc with white or coloured sectors adjoining black sectors is rotated. A black sector following a white one then appears to be covered by a light veil (the positive after-image). The speed of rotation can be so chosen that the last part of the black sector looks practically black. From the speed of rotation and the breadth of the black sector we can then calculate the time during which the positive after-image makes its influence felt. The figures found by Plateau in this were 0.35 sec. for white and yellow; 0.34 sec. for red and 0.32 sec. for blue.

These results, obtained in a manner analogous to our experiments with intermittent light, are in remarkably good agreement with the time of 333 msec. during which, in our opinion, the positive after-image favours the identification of



the homolateral squares projected with intervening pauses. This can hardly be a coincidence. After a pause of 333 msec. the positive after-image has presumably lost its influence; this is also compatible with the fact that in our experiments with a time of 333 msec. the fluctuations of brightness were so strong that the surrounding black appeared to pass over the squares from time to time — although this was not always the case.

A second problem is the fact that tunnel movement makes its appearance at a projection time of 500 msec., whereas the first sign of a completely normal apparent motion does not appear until a projection time of 1,667 msec. is reached. In both cases we are concerned with an apparent movement; this suggests the labile character of the original localization, which undergoes a change under the new conditions. Obviously it is the projection of the second object which, in this phase of 500 msec., influences the now labile localization of the first object or its after-image, and evokes an apparent motion which becomes reversed in the following 500 msec. and repeats itself rhythmically. It would appear that the projection time and intervening pauses must be lengthened to 1,667 msec. before (a) the time interval between homolateral projections has increased so far that the road is left open for identification with the object which is spatially separated but which follows immediately in time and (b) the influence of duration of projection has reached the height at which the spatial distance is overcome and identification of place results in movement. If the course of events is here correctly represented, it follows that with a shorter distance between the 2 objects a shorter projection time will be required to give a normal apparent motion. The significance of the duration of projection of the first object for the apparent motion had already been established by van der Waals and Roelofs at an earlier date.

In the third place we have the remarkable phenomenon of the after-image motion. This was always seen to appear after some time, a fact that is easily understandable when it is remembered that the intensity of the negative after-image increases progressively. Motion of the negative after-image cannot be expected until this has displaced the positive after-image, i.e. when the pause between 2 projections at the same place is longer than the duration of the positive after-image and long enough to

give the negative after-image time to reach a certain intensity. Under the conditions used this was the case with a pause of about 1,667 msec. Now, however, another question arises: why under these conditions does it continue to predominate over the normal apparent motion? It appears probable that the after-image is swallowed up by the intervening black, movement of which is seen at a quite early stage and which forms the moving background against which the apparent motion takes place. In normal apparent motion of the square, this is in the opposite direction to that of the intervening black, which at the same time constitutes the background.

There is yet another argument for the predominance of the after-image movement. A priori, both an object movement and an after-image movement may be expected. There is, however, a double tendency to identification: Identification of the objects separated in space excludes identification of the objects separated in time. The alternating perception of after-image motion and apparent motion shows that we are on the border line, where the 2 tendencies to identification prevail in turn. An identification of the homolateral objects projected at short intervals of time on the same place is quite compatible with an after-image movement. The more the intensity of the negative after-image increases, the more easily will the after-image motion occur and the less will be the obstacle to identification of the homolateral objects.

GROUP II. Distance from A to B 60 cm or  $13^{\circ} 1' 8''$  (see graph. I)

*Projection time  $\frac{1}{6}$  sec.* The phenomena were practically the same as with a distance of 80 cm between A and B. Flickering squares without apparent motion. When one square was fixated the other appeared to flicker more slowly and irregularly. When the intervening black was fixated, a black strip was seen to move to and fro; it was bounded by the haloes round the squares.

*Projection time  $\frac{1}{4}$  sec.* The squares seemed to rise somewhat as they became lighter and to move down again as they became darker (balance movement). Upon fixation of the middle of the space between the 2 squares, both squares were often seen swinging back and forwards with small excursions at the same time. This shows that apparent motion does not invariably

represent identification but points in the first place to an influence on localization; on both sides there is an alternating motion from R to L and from L to R, entirely analogous to the up and down movements; both objects move in the same direction without any sign of approach. In normal optimal apparent motion the identification is not produced by approach in space either, but by a yielding of the relative localization in favour of the labile egocentric localization and identification in time. We shall return to this point (see p. 169 and 173).

Here again the phenomenon chiefly noted is the movement of the intervening black. The to and fro movement of the squares is in the opposite direction to the black movement, but the former often seems slower and more irregular. There is, thus, a first trace of apparent motion but not yet any identification (weak apparent motion with simultaneity).

*Projection time  $\frac{1}{3}$  sec.* As the objects became lighter they moved slightly away from the middle and upwards as they became darker they went slightly towards the middle and downwards. There was thus, as with a projection time of  $\frac{1}{4}$  sec., a simultaneous to and fro oscillation of the squares in addition to the balance movement. The displacement of the intervening black, at first bounded by the haloes, extended gradually further so that it appeared as if the black passed in front of the squares from time to time.

*Projection time  $\frac{1}{2}$  sec.* As with the distance of 80 cm between A and B, the first sign of tunnel movement was now seen, but the movement of the intervening black still predominated.

*Projection time  $\frac{2}{3}$  sec.* Alternation of tunnel movement and black movement, the latter still predominating to some degree. Although the black seemed to push in front of the squares, it was not yet possible to speak of an after-image movement.

*Projection time 1 sec.* Here again alternation of tunnel movement and black movement, but the latter no longer predominated. During the transition from tunnel to black movement, the 2 motions could be seen simultaneously for an instant, taking place in opposite directions.

*Projection time  $\frac{5}{4}$  sec.* With this time all the forms of



apparent motion so far discussed could be seen in turn. The tunnel movement predominated. It was often accompanied by a simultaneous but oppositely directed black movement of the central field. Real after-image movement was seen on several occasions, the squares then apparently remaining motionless. Once or twice there was even a trace of normal apparent motion.

*Projection time  $5/3$  sec.* At first practically nothing but tunnel movement was seen, with only now and then a weak normal apparent movement, i.e. this was a transition between tunnel and normal apparent motion. Later the tunnel movement alternated with a true after-image movement, the squares now appearing to be motionless.

*Projection time  $5/2$  sec.* Tunnel movement merging into and alternating with clear normal apparent motion. Later after-image movement also. Sometimes the object movement and after-image movement were visible side by side in opposite directions; the object and the after-image approached and one passed in front of the other.

#### *Discussion of the phenomena of Group II*

As with a distance of 80 cm between A and B, we saw for the first time at a projection time of 333 msec. that the black now and then appeared to pass completely in front of the white square, while the identification of the homolateral projections could not be abolished completely until an interval of 500 msec. between these projections was reached. We are of the opinion that the duration of the positive after-image plays a part in this.

The balance movement, which actually has nothing to do with the kind of apparent motion studied here, also requires a moment's consideration. In other experiments we had also repeatedly noticed that when a light appears in darkness, this light seems to move from below upwards as it appears. In view of this continually recurring phenomenon we feel justified in assuming that an innervation to deorsumversion is evoked as a consequence of involuntarily occurring optomotor reflexes, so that an innervation to sursumvergence is necessary for fixation of the source of light, as a result of which the localization is displaced upwards.

With the very short projection times a simultaneous to and



from oscillation of the 2 squares was reported. In this case there is still no question of any identification of A and B. Nevertheless, the localization of A and B is reciprocally influenced by the appearance and disappearance of the 2 squares in turn (see above, p. 167 and 173). It is possible that this apparent oscillation is promoted as a contrast effect by the black movement of the central field. But this is certainly not the only cause. From previous investigations by van der Waals and Roelofs we know that such an apparent movement with simultaneity can also be perceived without apparent motion of black.

When considering the results of Group I we wondered whether a longer projection time of the preceding object could cause a displacement of the following object over a longer distance. In that case the tunnel movement would be replaced by normal apparent movement, even at shorter projection times, when the distance between the heterolateral objects was shorter. There is some evidence of this as with a distance of 60 cm a normal apparent motion is occasionally seen even with a projection time as short as  $\frac{5}{4}$  sec. Further, with a projection time of  $\frac{5}{4}$  sec. we saw the black movement of the central field replaced by an after-image movement. Black movement and after-image movement are not identical phenomena. In the after-image movement it seems as though a black square moves over the dark central field all the way from A to B and back again.

In the cases in which tunnel movement or normal apparent motion alternate with black movement or after-image movement, we might speak of a competition between time and space. A shorter pause between the homolateral projections will make the localizations of these projections more stable and thus oppose apparent motion, while with a longer pause the lability gives the apparent motion a free hand. We can say the same thing in a different way by pointing out that in normal apparent motion the resistance of the time interval against identification is the stronger; such motion will thus be more frequent as the time interval between the homolateral projections becomes greater. In black movement and after-image movement in the region between spatially separated objects, on the other hand, the resistance of the spatial interval between the heterolateral projections is the stronger; this will be especially the case when the time interval between the homolateral projections is shorter.

GROUP III. Distance between A and B 40 cm or  $9^{\circ} 8' 52''$  (see graph. I).

*Projection time  $\frac{1}{6}$  sec.* Chiefly 2 flickering squares. When one square was fixated, the peripheral square showed small horizontal apparent movements over a very short distance while its flickering was less noticeable. When the central field was fixated, both squares oscillated to and fro at the same time in a horizontal direction. Sometimes one had the impression that the distance between A and B did not always remain the same. Upon fixation of the central field we also saw the to and fro oscillation of a black strip bounded by the haloes of the 2 squares.

*Projection time  $\frac{1}{4}$  sec.* Flickering squares. Chiefly black movement of the intervening space, which was not confined to the central field but covered each of the squares in turn without eclipsing them completely. On fixation of the central field we sometimes again saw simultaneous oscillations of the 2 squares.

*Projection time  $\frac{1}{3}$  sec.* First we saw black movement bounded by the haloes, but soon there appeared black movement with covering of the objects in turn, without complete eclipse. Very slight motion of the objects, which seemed to come from centrally and below as they became lighter and to disappear towards the centre and downwards as they became darker. This means, thus, a simultaneous oscillation of the 2 squares.

*Projection time  $\frac{1}{2}$  sec.* Alternating tunnel and black movement but no after-image movement as yet. The black movement still predominated.

*Projection time  $\frac{2}{3}$  sec.* Marked predominance of tunnel movement, sometimes seen as apparent motion with a trace of simultaneity. Now and then after-image movement.

*Projection time 1 sec.* Alternation of tunnel and after-image movement. At the beginning of tunnel movement a black movement in the opposite direction was often visible. Once or twice a fairly distinct normal apparent motion was also seen. On the whole the tunnel movement predominated.

*Projection time  $\frac{5}{4}$  sec.* Alternation of tunnel and normal apparent motion, or a transition between them. Sometimes also a tunnel movement alternating with after-image movement.

*Projection time  $\frac{5}{3}$  sec.* Chiefly and primarily normal apparent motion. Later also after-image movement; sometimes transition of this into a tunnel movement with movement of black in the opposite direction.

*Projection time  $\frac{5}{2}$  sec.* Normal apparent motion. This was frequently optimal, but still often with a trace of tunnel movement; in particular when attention was paid to the black between the 2 haloes the object seemed to move under this intervening black.

#### *Discussion of the phenomena of Group III.*

Although the spatial distance between the 2 projections is now smaller, the first apparent motion with identification, in the form of tunnel movement, still does not appear until the time interval between 2 homolateral projections has been brought up to 500 msec. The shorter spatial distance between the heterolateral projections does, however, influence the after-image motion and normal apparent motion. These 2 movements are seen earlier, i.e. at a shorter time interval between the homolateral projections than was the case with a spatial interval of 60 or 80 cm. between the heterolateral objects. Identification of the 2 heterolateral objects has thus become easier. This is shown still more clearly at a projection time and time interval of 2.500 msec.; then the black after-image movement has completely disappeared, which means that identification of the homolateral objects is no longer taking place and we are dealing exclusively with an identification of heterolateral objects.

GROUP IV. Distance from A to B 20 cm or  $4^{\circ} 34' 52''$  (see graph. I)

*Projection time  $\frac{1}{6}$  sec.* Chiefly motion of the intervening black, which appears to pass repeatedly in front of the squares, but without eclipsing them completely, so that 2 squares surrounded by flickering haloes remain continually visible. Now and then the squares appear to move simultaneously somewhat, the distance between them also appearing to change slightly.

*Projection time  $\frac{1}{4}$  sec.* At first only movement in a vertical direction, upwards with increase of brightness and downwards with decrease of brightness. At the same time, black movement which seems to undergo transition into after-image movement.

Owing to the very short distance between the 2 projections it is difficult to make out whether a true after-image movement is present. Now and then a simultaneous movement of the 2 squares is seen.

*Projection time  $\frac{1}{3}$  sec.* Alternation of tunnel movement and after-image movement.

*Projection time  $\frac{1}{2}$  sec.* Very distinct normal apparent motion with a trace of simultaneity. Occasionally also still after-image movement, especially upon fixation of the intervening black.

*Projection time  $\frac{2}{3}$  sec.* Very distinct normal apparent motion with a trace of simultaneity. Still an occasional after-image movement.

*Projection times 1 sec. to  $\frac{5}{2}$  sec.* Optimal apparent motion.

#### *Discussion of the phenomena of Group IV*

It is not surprising that with this short distance between the 2 projections the apparent motion is manifested earlier than with the longer distances. But there are certain phenomena which demand attention. In the first place, the first sign of apparent motion of the objects (tunnel movement) was seen with a time interval of only 333 msec. between the homolateral projections. From the preceding results, in agreement with information obtained from the literature, we felt justified in concluding that at this time the positive after-image has not yet disappeared. It is therefore difficult to assume that abolition of the identification of the homolateral projections has already occurred. The most obvious explanation would be that, owing to the very short distance between the 2 projections a reciprocal influence on the localization is at work, still without there being any direct identification of the heterolateral squares. Another point is that as a consequence of the 2 projections being so very close together the rest of the positive after-image is suppressed by simultaneous contrast, this leaving the coast clear for an earlier appearance of a negative after-image. Even with a pause of only 250 msec. a trace of after-image movement was seen. This is a movement of the negative after-image; in the pause of 250 msec. it must, thus, already have been possible for a negative after-image to form in succession to the positive



after-image. We must admit that with a projection time of 250 msec. it was difficult to distinguish between black movement and after-image movement, but with a projection time of 333 msec. there was undoubtedly an impression of movement of the negative after-image. With a projection time of 1 sec. or longer the rivalry between time and space with respect to identification is settled; the time interval then offers stronger resistance than the spatial interval and an optimal apparent motion is continuously present.

GROUP V: Distance A—B 10 cm or  $2^{\circ} 17' 30''$  (see graph. I)

*Projection time  $\frac{1}{6}$  sec.* Often a very distinct apparent motion, although with doubling so that it would perhaps be better to speak of simultaneous motion. This simultaneous motion became particularly evident when one fixated 20 cm or more to one side. Upon fixation of the squares, both of them were sometimes seen to keep still while a black patch seemed to waver at the boundary line between the 2 squares seen as a light strip. This light strip is probably due to the fact that the 2 squares are not seen side by side but seem to overlap somewhat. The black patch does not cover the squares completely. It is presumably due to the strong contrast between the place of the square which has just appeared and that of the square which has just disappeared, i.e. the continual alternation of the place of light and darkness. This is, thus, not a true after-image movement.

*Projection time  $\frac{1}{4}$  sec.* Optimal apparent motion, occasionally replaced by movement of black or after-image movement. Continual black movement between the haloes above and below the objects, this black movement being in the opposite direction to the apparent movement.

*Projection time  $\frac{1}{3}$  sec.* Optimal apparent motion with a trace of simultaneity.

*Projection times  $\frac{1}{2}$  to  $\frac{5}{2}$  sec.* Optimal apparent motion.

#### *Discussion of the phenomena of Group V.*

Since the 2 squares are objectively just touching, no tunnel movement is to be expected here. As already explained, the movement phenomena with a projection time of  $\frac{1}{6}$  sec. are of

such a kind as to suggest a simultaneous movement rather than a normal apparent movement. This simultaneous movement is the same as that seen earlier, when it was also distinguished from apparent motion in the more restricted sense (see p. 167 and 169).

A remarkable fact is that this movement is more distinct when one fixates 20 cm or more to one side, probably because both objects are then equally passive. This last observation is reminiscent of phenomena observed by Hylkema in his investigation of the fusion frequency in different parts of the field of vision. He found the fusion frequency for small test fields of diameter  $1.5-3^\circ$  to be greatest 10-25 degrees nasal to the fovea and 45 degrees temporal to the fovea. This fusion frequency was lowest in the fovea and the immediate neighbourhood of the blind spot. A higher fusion frequency indicates a more rapid loss of intensity of the positive after-image and such a more rapid loss of intensity of the positive-after image will make the localization of the homolateral projections less stable and thus render an apparent motion earlier possible. An indubitable normal apparent motion was found with a projection time of  $\frac{1}{4}$  sec. This also raises certain problems. With a larger distance between the 2 projections there was still no sign of a normal apparent motion with a projection time and pause of 250 msec. The projected squares did indeed appear to flicker, but their brightness did not decrease so far that they became invisible. How can we now account for this in conjunction with a normal apparent motion? It might be suggested that the positive after-image disappears earlier in consequence of the strong contrast with the following projection adjacent to it. It seems more likely, however, that the positive after-image actually does not disappear earlier but is localized elsewhere, i.e. in the direction of the subsequent projection. This conjecture is supported by the fact that in apparent motion over this short distance the brightness of the square remains perfectly constant. It is also in good agreement with the suggestions and explanations presented in the foregoing (p. 167, 169 and 173).

Nevertheless, we do not wish to exclude entirely the possibility that the positive after-image does disappear sooner owing to the strong contrast with the adjacent following projection. Both possibilities might be present at the same time. This might be

supported by the fact that after-image movement (of the negative after-image) can also be seen alternating with the normal apparent motion. As soon as this happens the homolateral projections are identified, but with a moment of invisibility. The intensity of the positive after-image must have decreased to a greater degree than with a larger distance between the 2 projections. As a matter of fact the occurrence of the negative after-image with the short pause of 250 msec. is in itself a problem, as we have already pointed out. In this connection it should be remembered that the negative after-image may even develop during the existence of the positive after-image, becoming manifest only after the processes responsible for the positive after-image have become weaker.

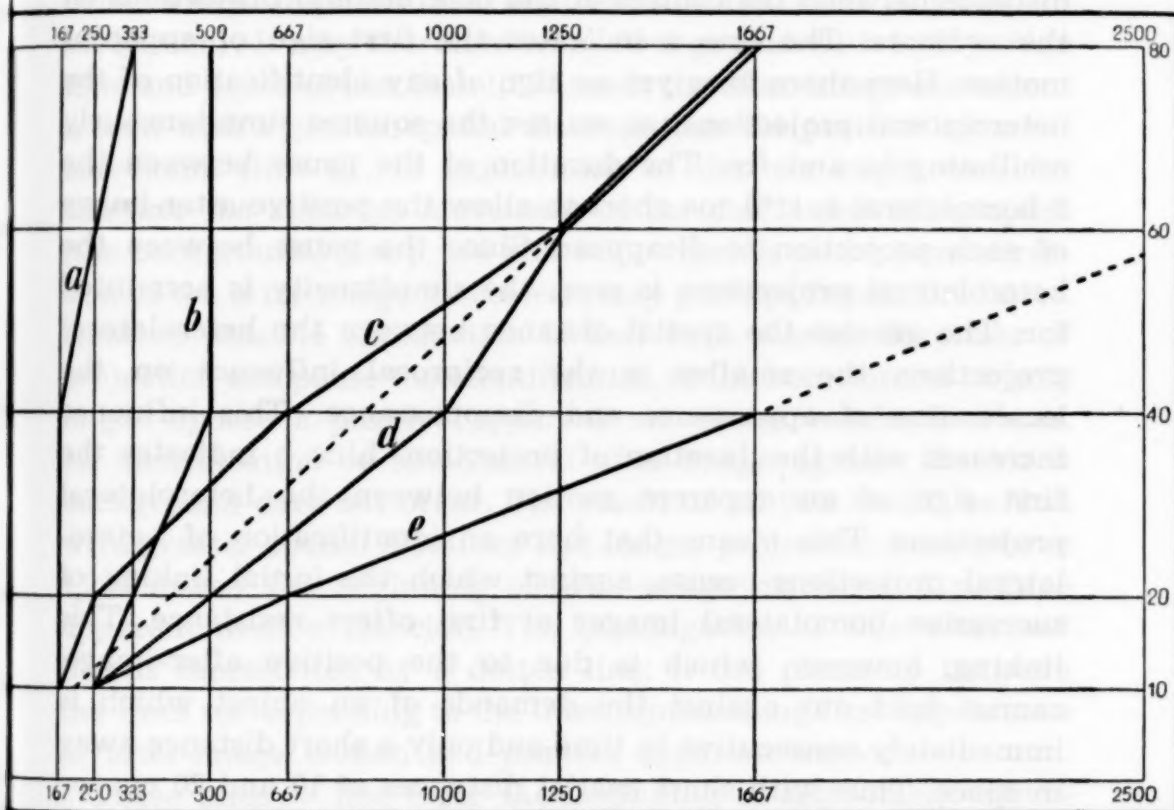


Fig. 2.

*Discussion of the results as a whole (see graph. 2)*

After discussing the phenomena of each group separately we can now, in a comprehensive survey, pay attention to the various phenomena of motion and their appearance and disappearance,

in so far as this is influenced by the given conditions of time and space.

In the foregoing we have dealt with the results of the investigation in separate groups, each of which corresponded to a single spatial distance between A and B. We could also have arranged the results in such a way that each group corresponded to a certain projection time and duration of the pause between homolateral projections, thus viewing the phenomena from a different angle. But even without this we are of the opinion that the various aspects of the observed phenomena receive full justice in the following survey.

In graph. II the projection time or time interval between 2 homolateral projections is plotted on the abscissa and the spatial distance between the centres of the heterolateral projections on the ordinate. The line a indicates the first sign of apparent motion. Here there is as yet no sign of any identification of the heterolateral projections, as we see the squares simultaneously oscillating to and fro. The duration of the pause between the 2 homolateral is still too short to allow the positive after-image of each projection to disappear. Since the pause between the heterolateral projections is zero, the simultaneity is accounted for. The greater the spatial distance between the heterolateral projections the smaller is the reciprocal influence on the localization of appearance and disappearance. This influence increases with the duration of projection. Line b indicates the first sign of an apparent motion between the heterolateral projections. This means that here an identification of heterolateral projections occurs, against which the initial linking of successive homolateral images at first offers resistance. This linking, however, which is due to the positive after-image, cannot hold out against the demands of an object which is immediately consecutive in time and only a short distance away in space. Thus with short spatial distances of 10 and 20 cm we already see signs of tunnel and true apparent motions with projection times that are still relatively short — 250 and 333 msec. But although even with these short projection times the duration of the positive after-image is too short to bridge over the pauses and hence leaves a gap in the resistance which it offers, the opportunity thus presented is still insufficient to permit apparent motion when the spatial distance between the heterolateral



projections is *more* than 20 cm. When the projection time and pause are increased to 500 msec. or longer, the coast is now clear for identification with heterolateral projections, independently of the spatial distance whether this be 40, 60 or 80 cm.

Curve c of graph. II gives the first appearance of a movement of the negative after-image, with various distances between A and B and progressive increase of the projection time and time interval between homolateral projections. We cannot remember ever having found mention of this after-image motion in the literature; the evocation of this form of apparent motion is presumably due to the special conditions of our investigation. The positive image and the negative after-image in our investigation are equal in motor condition (the same points on the retina from which impulses are emitted and the same rhythm, of succession). It may therefore seem strange that with a projection time and pause of less than 1,250 msec. the after-image movement is seen with a greater spatial distance than the positive apparent movement (line d). In explanation of this we might consider whether the apparent movement of the dark intervening space (the background) (see p. 166 and 169), which is in the same direction as the after-image movement, may perhaps favour the latter. Another possible explanation is that with shorter durations of projection and pause the identification of the homolateral objects might oppose a positive apparent movement, while, however, an identification of homolateral after-image against the black background does not occur. The result of all this would be that with greater spatial distances and longer projection time (above 1,250 msec.) the beginning of after-image motion and positive apparent motion coincide. The prolongation of this combined line is represented by a dotted line; it lies plausibly between the lines corresponding to the lines representing first appearance of after-image motion and positive apparent motion.

These lines — line c for the beginning of after-image movement and line d for the beginning of normal apparent motion — are not far apart. The appearance of a negative after-image seems to herald the labile phase in which the spatially separated heterolateral square can exert its influence on the localization of this residual negative image. At the transition from after-image motion to positive apparent motion we have the narrow no-man's-land in which temporal and spatial resistances just

balance each other. If we imagine a curve lying between *c* and *d* this will coincide more or less with the straight dotted line. At every point on this line the ratio of the time interval between homolateral projection (in msec.) to the spatial distance between heterolateral projections (in cm) is approximately 20.8. This means that increase of the time interval raises the resistance to identification of homolateral projections in the same way and to the same degree as increase of the spatial distance raises the resistance to identification of heterolateral projections.

In the registration of our observations we have sometimes made use of the term 'tunnel movement'; the phenomena thus indicated require some comment — the more so as the occurrence of this tunnel movement under certain conditions seems at some points to be incompatible with the explanations offered in the foregoing. These discrepancies prove, however, to be unreal when we consider the meanings of the terms used. It then appears that the phenomenon of so-called tunnel movement involves phenomenologically movement but not the identification which the term suggests; it lacks the convincing direct character of identification of a concealed separateness which belongs to true apparent motion in the strict sense of the term. While in normal apparent motion only one square is perceived, which has escaped from the regulation of egocentric localization and is governed only by the varying interplay of reciprocal influencing of the impressions received, in the so-called tunnel movement there is no such directly evident identification. At the most there is here a secondary identification in that we tend to regard something that disappears in one direction and something that appears from the other direction as one and the same.

Although at first it appeared to us as though the 'tunnel movement' was a preliminary stage of true apparent motion and that only a gradual and not an essential difference existed between the 2, an essential difference between them is revealed by the facts reported above and by more profound consideration. In the former case we are still dealing with a phenomenon of rhythmically-changing localization only, while in the latter we are furthermore dealing with an identification. The tunnel movement has the further peculiar feature that we can distinguish 3 phases in it: disappearance of the 1st object,

invisibility of both objects and appearance of the 2nd object, and also that the momentary invisibility of the objects is favoured by concentrating attention on the dark central field. This makes it probable that there is at the same time a certain degree of suppression, which belongs to the class of competitive phenomena. This would also account for the fact that it is precisely with the larger spatial distances that the tunnel movement persists so much longer.

Finally we have also plotted a line *e* on graph. II. This shows the end of the after-image movement. Probably the tendency to identification of the homolateral projections has then become so small that it no longer opposes any resistance to the apparent motion. But then also the stronger stimuli of the luminous squares will be more likely to give rise to normal apparent motion than the weak stimuli of the dark after-images to give rise to after-image motion. Curve *e* is also practically a straight line and when we prolong this we are able to understand that with a projection time of 2,500 msec. and a spatial distance of 60 cm we still sometimes see an after-image movement. The curve shows that with twice the distance between A and B the projection time of the homolateral projections must be 2.5—2.6 times as large if a normal apparent motion is to be ensured. It is obvious that this relationship is valid only within certain limits; the distance over which the apparent movement takes place cannot be increased indefinitely by prolonging the projection time of the objects. Further, we see here again that a time interval *a* times as long between the homolateral projections increases the resistance to identification to roughly the same degree as a spatial distance *a* times as large between the heterolateral projections.

#### SUMMARY

A white square was projected intermittently on a screen, with durations and pauses of the same length. At a certain distance to the side of this a 2nd white square was projected, alternating accurately with the first. In a first series of experiments the projection time and pause was progressively increased from a very short (167 msec.) to a considerably longer interval (2,500 msec.).



In the following experiments the distance between the 2 projections was varied from 10 to 20, 40, 60 and 80 cm.

As was to be expected, when the projection time and pause was very short the successive projections were identified and fused, so that 2 objects separated in space were perceived.

With slower alternation, i.e. longer projection and pause, the first square was identified with the immediately consecutive square situated at a lateral distance from it, as a result of which apparent motion was seen.

Special phenomena, e.g. the after-image movement, which appeared in this investigation are discussed.

In a similar way to that in which lengthening of projection time and pause hinders homolateral identification, an increase of the spatial distance creates a resistance to heterolateral identification.

In this way it was found that great similarity exists in the ways in which time intervals and spatial intervals can influence the identification of 2 successive projections.

This demonstrates once more how intimate is the connection established between 'time' and 'space' by the personality which co-ordinates and graduates both and creates and experiences them in the continuity of its own change and movement.

#### ZUSAMMENFASSUNG

Ein weisses Viereck wurde derartig intermittierend auf einem Schirm projiziert, dass die Projektionsdauer und die Unterbrechung immer von genau gleicher Grösze waren. In einiger Entfernung neben diesem Viereck wurde ein zweites, ganz gleiches Viereck projiziert, dessen Projektionsdauer sehr präzise mit der des ersten Viereckes alternierte.

In einer ersten Versuchsreihe wurden Projektionsdauer und Unterbrechung allmählich geändert, von ganz kurz (167 msec.) bis bedeutend länger (2500 msec.). In einer folgenden Versuchsreihe wurde die Entfernung zwischen den beiden Vierecken geändert, sodass diese nacheinander 10, 20, 40, 60 und 80 cm betrug.

Wie zu erwarten war, wurden bei ganz kurzen Projektionsdauern und Unterbrechungen die an derselben Stelle aufeinander folgenden Projektionen identifiziert und fusioniert, sodass zwei räumlich getrennte Vierecke wahrgenommen wurden.



Beim trägeren Alternieren, d.i. bei längerer Projektionsdauer und längerer Unterbrechung, wurde das eine Viereck mit dem zeitlich gleich darauf folgenden anderen Viereck identifiziert, wodurch eine Scheinbewegung sich manifestierte.

Besondere Erscheinungen, unter anderen die Nachbildbewegung, die sich bei diesen Untersuchungen darboten, werden näher erörtert.

In ähnlicher Weise, in welcher Verlängerung von Projektionsdauer und Unterbrechung eine homolaterale Identifizierung erschwert, schaffte eine Vergrößerung der räumlichen Entfernung einen Widerstand gegen die heterolaterale Identifizierung.

Also manifestierte sich eine grosse Übereinstimmung in der Weise, in welcher zeitliche und räumliche Abstände die Identifizierung von zwei nacheinander folgenden Projektionen beeinflussen können.

Wiederum kommt hierin zum Ausdruck die innige Verwandtschaft zwischen „Zeit“ und „Raum“, koordiniert und graduiert durch unsere Persönlichkeit, welche die beiden in der Kontinuität von Selbstveränderung und Bewegung nach eigenem Maszstabe schafft und erlebt.



## IN MEMORIAM GUSTAV KAFKA

VON

G. RÉVÉSZ

50 Jahren dürfte es her sein, dass ich in der Weenderstrasse zu Göttingen vor einer Buchhandlung stand. Plötzlich blieb ein langer, elegant gekleideter junger Mann neben mir stehen und betrachtete ebenso die Neuerscheinungen der Psychologie. Ich fragte ihn, ob auch er Psychologie studiere; und als er meine Frage mit einem kurzen Satz beantwortete, wusste ich sofort, dass er ein Wiener sei. Von diesem Augenblick an datiert unsere Freundschaft, die bis zu seinem Tode dauerte.

Kafka hat später Göttingen verlassen, um bei W. Wundt 1904 in Leipzig zu promovieren. Dann ging er nach München, wo er 1910 Privatdozent und 1915 o.a. Professor wurde. In München gehörte er zu dem intimsten Kreise des auf seine Schüler so starken Einfluss ausübenden Theodor Lipps.

Den ersten Weltkrieg hat Kafka als Reserverittmeister des oesterreichischen Heeres auf dem europäischen und kleinasiatischen Kriegsschauplatz mitgemacht und erst im letzten Jahr des Krieges wurde er beauftragt, ein Heerespsychologisches Institut in Wien zu errichten. Zur gleichen Zeit gründete auch ich eine psychologische Eignungsprüfstelle für das ungarische Heer in Budapest, was uns Gelegenheit bot, die Eignungsprüfungsmethoden für Spezialtruppen gemeinsam auszuarbeiten und anzuwenden.

Nach Ablauf des Krieges ist Kafka zu seiner Familie nach München zurückgekehrt. Im Jahre 1923 erhielt er eine Berufung an die Technische Hochschule in Dresden als ordentlicher Professor für Philosophie, Psychologie und Pädagogik.

Ganz im Beginn des Hitlerregimes kamen die führenden Psychologen Deutschlands in Berlin zusammen, um zu der neuen politischen Lage Stellung zu nehmen. Einige von ihnen nahmen den Standpunkt ein, die jüdischen Professoren müssten zum Wohl der Nation von jeder Lehrtätigkeit ausgeschlossen werden, und so mussten die bei der Sitzung anwesenden

William Stern und David Katz erleben, dass sie von ihren nationalsozialistischen Kollegen wegen ihres Judentums für jede wissenschaftliche Lehrtätigkeit als ungeeignet und unwürdig erklärt wurden. Zwar waren nicht alle anwesenden Psychologen mit der Haltung der Nationalsozialisten einverstanden, aber — wie ich erfuhr — war Gustav Kafka der einzige, der sich mit seiner ganzen moralischen Kraft diesem Entschluss entgensetzte, was allerdings auf die weitere Entwicklung der Dinge keinen Einfluss ausübte, aber auf Kafka's Charakter ein bezeichnendes Licht wirft.

Infolge der damaligen Aufregungen ist Kafka schwer krank geworden, was zur Folge hatte, dass es ihm gelang, sich von der Hochschule gänzlich zurückzuziehen und als Privatmann seine wissenschaftliche Arbeit in grosser Einsamkeit fortzusetzen.

Nach Unterwerfung der deutschen Waffen wurde Dresden bombardiert. Nur mit grösster Not gelang es Kafka und seiner Familie aus der brennenden Stadt zu flüchten. Ihr ganzes Haus und die Einrichtung der Wohnung, die so viel schöne Dinge aus der Wiener Zeit besass, nebst seiner wertvollen Bibliothek sind vollständig verbrannt. Kafka wollte nicht in Dresden bleiben, und als die Russen die Stadt besetzten, bemühte er sich nach Bayern zurückzukommen. Mit Freude hat er 1947 die Berufung der Universität Würzburg angenommen und mit persönlichen Opfern und grösster Energie hat er das Psychologische Institut wieder aufgebaut und die psychologische Forschung in Gang gesetzt. Bescheidene Unterkunft fand er mit Frau und Tochter in einem Bau der Universitätskliniken. Erst nach seinem Emeritatus gelang es Frau Kafka, die ihr ganzes Leben ihrem Manne in vorbildlicher Weise opferte, in der Nähe von Würzburg in ländlicher Umgebung ein nettes Heim zu finden, wo mein armer Freund Kafka nach so vielen schweren Jahren wieder Ruhe und freie Natur hätte geniessen können. Leider dauerte das idyllische Leben nur einige Monate, denn am 12. Februar 1953 ist er infolge eines Herzschlages plötzlich verschieden.

Kafka hinterliess seine Frau und seine junge verheiratete Tochter, ferner zwei Söhne aus erster Ehe mit einer ungemein lieben Wienerin, die bald nach dem ersten Weltkrieg verschied. Der älteste Sohn, der wegen antinationalsozialistischer Arbeit hier in Holland durch die Gestapo arrestiert wurde und während des ganzen Krieges im Gefängnis sass, bekleidet eine Stelle in



einem Verlag in Graz und will sich als Jurist an der Universität habilitieren; der etwas jüngere Sohn, der ebenfalls unter den Nazis verhaftet war, ist Revierförster in Oberbayern.

Der bestechendste Zug der wissenschaftlichen Persönlichkeit Kafka's war seine weitreichende allgemeine Bildung und sein grosses philosophisches Interesse. Er gehörte zu der äusserst kleinen Gruppe der Psychologen, die mit gleicher Hingabe philosophische *und* psychologische Probleme behandelte. In allen seinen psychologischen Schriften fühlt der Leser seine philosophische Einstellung in der Problematik wie in der Methode. Alles, was er veröffentlicht hat, zeigt Gelehrsamkeit und Scharfsinn; des Ganze ist stets logisch gegliedert, jedes Einzelne sich wieder zu einem abgerundeten Ganzen erweiternd.

In der von ihm herausgegebenen „Geschichte der Philosophie in Einzeldarstellungen“ hat Kafka von 1923 an vier sehr schöne Bücher veröffentlicht, nämlich die „Vorsokrater“, „Platon, Sokrates und der sokratische Kreis“, „Aristoteles“ und gemeinsam mit dem Philosophen Eibl den „Ausgang der Antike und das Erwachen einer neuen Zeit“. Später sind von ihm erschienen: „Geschichtsphilosophie der Philosophiegeschichte“ (1933), „Naturgesetze, Freiheit und Wunder“ (1940), „Was sind Rassen?“ (1947) und „Freiheit und Anarchie“ (1947).

Von seinen psychologischen Werken will ich besonders hervorheben die „Tierpsychologie auf experimenteller und ethnologischer Grundlage“ (1914) und seine „Tierpsychologie“ im „Handbuch für vergleichende Psychologie“ (1923), die bei den Psychologen und Biologen grosse Anerkennung fanden. In der von mir herausgegebenen internationalen Zeitschrift „Acta Psychologica“, deren Mitherausgeber er war, hat er drei sehr wichtige Abhandlungen publiziert, nämlich „Grundsätzliches zur Ausdruckspsychologie“ (1937), „Über das Erlebnis des Lebensalters“ (1949) und „Uraffekte“ (1952). In das anlässlich meines 70. Lebensjahres herausgegebene Spezialheft der Acta Psychologica (1950) schrieb er einen durch Wärme und Verständnis erleuchteten Aufsatz unter dem Titel „Erinnerungen eines Jugendfreundes“.

Kafka war einer der bedeutendsten Repräsentanten der Psychologie in den deutsch sprechenden Ländern. Er wurde wegen seinen hohen moralischen Eigenschaften, seiner lebenswürdigen Persönlichkeit und seinen wissenschaftlichen Ernst überall

geschätzt. Er bekleidete zahlreiche Ehrenämter in Deutschland und im Ausland. Durch seine Unabhängigkeit des Denkens und nicht weniger infolge der Mannigfaltigkeit seiner Forschungsobjekte hat er eine besondere Stellung unter den Psychologen Deutschlands erobert.

Unter meinen Kollegen stand Kafka mir menschlich so nahe, dass ich es als ein inneres Bedürfnis fühlte, das Lebensschicksal dieses würdigen, bescheidenen, weisen, tief religiös bewegten Mannes, der sich für die höchsten Ziele der Menschheit immer mutig, mit seiner ganzen Persönlichkeit einsetzte, niederzuschreiben. Die Psychologen verloren durch den Tod Gustav Kafka's einen ihrer Führer und Tröster; ich, meinen innigsten und stets verständnisvollen Freund.

#### IN MEMORIAM GUSTAV KAFKA

Fifty years may have passed since I was standing in front of a bookshop in Weenderstrasse, Göttingen, when suddenly an elegantly dressed tall young man stopped near me and also regarded the latest publications on psychology. I asked him if he also studied psychology, and when he answered my question with a short sentence I instantly knew that he was from Vienna. It was at that moment that our friendship began, and it has lasted until his death.

Some time later, Kafka left Göttingen in order to take his doctor's degree with Professor W. Wundt in Leipzig in 1904. From there he went to Munich where he was promoted Privatdozent in 1910, and assistant professor in 1915. In Munich he was a member of the most intimate circle around Theodor Lipps, whose influence on his disciples was very great.

During the First World War, Kafka served as a Captain of Cavalry in the reserve of the Austrian Army in the European theatre of war and that of Asia Minor, and not until the last year of the war was he commissioned with the erection, in Vienna, of an Institute for Military Psychology. At the same time I also founded an institute for psychological tests in the Hungarian Army at Budapest. Thus we had an opportunity to develop together testing methods for special troops and to apply them.

At the end of the war, Kafka returned to his family at Munich, and in 1923 he was called to the Technische Hochschule of Dresden as a full professor for Philosophy, Psychology, and Pedagogy.

At the very outset of the Hitler régime, the leading psychologists of Germany assembled in Berlin in order to adopt an attitude to the new political situation. A few of them took the view that, for the benefit of the nation, the Jewish professors should be excluded from all teaching activity. Here, the two professors William Stern and David Katz, who were present at the assembly, suddenly, on account of their being Jewish, heard their national-socialist colleagues declare them unfit for and unworthy of scien-

tific teaching. Not all of the psychologists present, it is true, shared the attitude of the National-Socialists. However, as I learned later, Kafka was the only one to oppose the resolution with his full moral strength. We know that the course of events was not changed through Kafka's conduct, but it throws a revealing light on his character.

In consequence of the emotional strain of those days Kafka fell seriously ill. Thus he succeeded in withdrawing completely from the University in order to continue his research work in great loneliness as a private individual.

In the course of the German defeat, Dresden was raided from the air, and it was only with the greatest difficulty that Kafka and his family succeeded in fleeing from the burning town. Their whole house was burnt down with its furniture that, beside his precious library, contained so many beautiful objects from his Viennese days. Kafka did not want to remain in Dresden, and when the Russians occupied the town he endeavoured to return to Bavaria. It was with great joy that he accepted the invitation to come to the University of Würzburg in 1947, where he rebuilt the Institute of Psychology with great energy and personal sacrifice and set psychological research in motion again. With his wife and daughter he was given most modest lodgings in one building of the university clinics, and not until after his retirement from the university, Mrs. Kafka, who, in an exemplary fashion, had shunned no sacrifice for her husband, succeeded in finding a pleasant home in the rural surroundings of Würzburg, where after so many years of hardship my poor friend Kafka could have enjoyed Nature in peace and quiet again. Unfortunately, that idyllic life was to last but a few months, for he passed away suddenly on February 12th, 1953, as a result of heart failure.

Kafka left behind his wife and his young married daughter, as well as two sons by his first wife who was a most charming Viennese lady who died soon after the First World War. His eldest son who, owing to anti-national-socialistic activities, had been arrested by the Gestapo in Holland and was detained in prison throughout the war, is employed at a publisher's in Graz and intends to be a professor for law at the university. The somewhat younger son, who was likewise arrested under the Nazis, holds a position as quarter-ranger in Upper-Bavaria.

The most engaging trait of Kafka's scientific personality was his wide general education and his great interest in philosophy. He belonged to the very small group of psychologists who are in like manner devoted to the treatment of philosophical and psychological problems. The reader feels Kafka's philosophical attitude in all his writings, in the problems as well as in the method. All his publications show both erudition and perspicacity. Everything is logically arranged, with each detail extending into another absolute unit.

In the "Geschichte der Philosophie in Einzeldarstellungen" (History of Philosophy, a Symposium), which was edited by Kafka from 1923 onward, he has published four valuable volumes: "Die Vorsokratiker" (The Pre-Socratics), "Platon, Sokrates und der sokratische Kreis" (Plato, Socrates and the Socratic Circle), "Aristoteles" (Aristotle), and in co-operation with

Professor Eibl, the philosopher: "Der Ausgang der Antike und das Erwachen einer neuen Zeit" (The End of Antiquity and the Dawn of a New Era). Later publications are: "Geschichtsphilosophie der Philosophiegeschichte" (1933) (A Philosophical History of the History of Philosophy), "Naturgesetz, Freiheit und Wunder" (1940) (Natural Laws, Freedom and the Marvellous), "What are Races?" (1947) and "Freiheit und Anarchie" (1947) (Freedom and Anarchy).

Of his psychological works, I want to lay particular stress on his "Tierpsychologie auf experimenteller und ethnologischer Grundlage" (1914) (Animal Psychology, on an experimental and ethnological basis) and his "Tierpsychologie" (Animal Psychology) which appeared in the Handbook of Comparative Psychology (1923) and which met with great approval with psychologists and biologists. Moreover, Kafka has published three important treatises in the International periodical "Acta Psychologica", edited by myself and of which he was a co-editor: "Grundsätzliches zur Ausdruckspsychologie" (1937) (The Essentials of the Psychology of Expression), "Über das Erlebnis des Lebensalters" (1949) (The Experience of Human Age), and "Uraffekte" (1950) (Primeval Affections). In the special edition of the "Acta Psychologica", which was published on the occasion of my 70th birthday (1950), Kafka wrote an essay entitled "Memories of a Friend of my Youth", which was inspired by warmth and understanding.

Kafka was one of the most outstanding representatives of psychologists in German speaking countries. Owing to his moral qualities, his charming personality, and his scientific integrity, he was held in high esteem everywhere. He held numerous honorary offices both in Germany and abroad. His independent thinking and, last but not least, the variety of objects of his research work have gained him a particular position among the psychologists in Germany.

Of all my colleagues, there was such close human relationship between Kafka and myself that I feel urged to write down the vicissitudes of life of this worthy, modest, wise, and deeply religious man who never lacked the courage to bring to bear his whole personality when the highest goals of mankind were at stake. Through Kafka's death psychologists have lost one of their consoling guides; I have lost my most intimate and always understanding friend.



*Department of Philosophy, University, Oxford*

## THINKING

BY

Prof. GILBERT RYLE

I want to ask two questions. The first question is Why has so little emerged from the painstaking investigations of psychologists in the theory of thinking? The second question, which undercuts the first, is What have these investigators supposed that they were looking for?

(1) Consideration of the problems tackled e.g. by the Würzburg School and its successors makes a philosopher like myself feel qualms of professional guilt. Again and again we find psychologists trying to observe, measure and describe just those ingredients or basic components of thinking operations which logicians and epistemologists have solemnly assured them must, on a priori grounds, be there. For example, some of our philosophical great-grandfathers declared that thoughts consist of ideas, variously originated and variously concocted. These ideas were then identified (since nothing else could be found with which to identify them), with mental images. Dutifully the experimental psychologists began to enquire how these ideas or images constitute thinking, or, when doubts arose, what their rôles in thinking are if they do not entirely constitute it, and what other ingredients can be detected filling in the gaps.

Our philosophical grandfathers, for excellent philosophical reasons, switched the brunt of their reflections from terms to propositions, from ideas to judgements; and dutifully the experimental psychologists laboured to observe, measure and describe the mental acts or processes of judging. Logicians and epistemologists have debated the notions of abstraction and generalisation; and the researchers dutifully set to work to isolate, under laboratory conditions, these officially sponsored acts or processes of abstracting and generalising. The unpalatable truth is that we philosophers have told epistemological fables, and the experimental psychologist has dutifully essayed the natural history of our fabulous monsters. What seems to have happened is this.

The logicians and epistemologists were in fact trying to give functional descriptions of the various kinds of elements into which constructed theories were analysable. Their eyes were on the terms, the connectives, the phrases, the propositions and the arguments of which published theories consist. But the descriptive apparatus that they employed for this end was, for some 300 years, a predominantly Cartesian or Lockean apparatus. They were trying to describe the functions of bits of theoretical discourse; but their descriptions were couched as narratives of introspectible mental processes. Thus they decoyed the experimental psychologists into the profitless enterprise of trying to provide systematised information about these mythical introspectibles — or, later on, the equally mythical unintrospectibles.

(2) An odd feature of the situation is this. Naturally the psychologist knows as well as the schoolboy knows, what thinking is. He couldn't *not* know, any more than the footballer couldn't *not* know what football is. He has practised it all his life, he got a lot of training in it at school, he has taught and examined other people, he has communicated his thoughts and followed the thoughts of others. He knows what it is like to be stumped, befogged, weary, puzzled and successful; what it is like to keep or lose the thread, to go round in circles, to give problems a rest and so on and so on. What, then, does he *not* know that he still wants to find out? For example, Watt and his subjects knew, before the experiments began, that trying to think out the answer to a question is not of a piece with saying 'Tweedledee' after someone else has said 'Tweedledum'. For saying 'Tweedledee' in this way is not giving the answer to a question; nor does saying it issue from wondering what is the right thing to say. In short association is notoriously not pondering, just as slithering is notoriously not marching. Yet a long series of patient experiments was supposed to be required to show that thinking cannot be dissected into associations.

Again, some experimentalists claim to have established on a firm experimental basis the generalisation that thinking is intimately connected with tasks or problems. But how did they conduct their experiments? By setting to their subjects tasks or problems which the experimenters already knew could be solved by thinking and could not be answered without thinking. It is as if someone should claim to have established inductively that

footballers try to score goals, and to have established this by getting them to play some games of football. If they did not know what football was, how could they set their subjects to engage in specimen games of it? It should be noticed that these experimentalists took care not to ask their subjects to tell how many chimney pots there are in London, or whether Shakespeare ever had mumps. They and their subjects were too well aware that such questions cannot be answered just by thinking. They did not discover the problem-solving nature of thinking by experimental questioning; they selected soluble questions for their subjects because they already knew that trying to solve such problems is thinking. To put it over crudely, it is a fact not of psychology but of grammar that the verbs 'wonder' and 'consider' are followed by indirect questions. Why then pretend to have to establish it experimentally — and by a method of interrogation which shows that it was known from the start anyway? Why confuse conceptual with empirical questions?

(3) I think that part, but only part of the tendency of logicians and epistemologists, and, after them, of psychologists to assume that there must exist some isolable and describable ingredients of thinking comes from this origin. If you ask me of what basic movements rowing or jumping consists, I can give you the answer, or at least I know how to find out the right answer to give. It is natural to suppose, then, that 'thinking' stands also for a specific process or activity, composed in various ways out of some common, recurrent, elements. Such a supposition is encouraged by the age-old dogma that our mental life is subdivided into three distinguishable strands or strata, Cognition, Conation and Feeling. This tripartite dogma itself suggests analogies from Chemistry. But if I asked you to tell me the basic elements of which *working* consists, or of which *gardening* consists or of which *housekeeping* consists, you would be quick to see the trap I was laying for you. You would say, quite rightly, that words like 'gardening', 'working' and 'house-keeping' cover a great number of widely different things. Two men may both spend their leisure hours in gardening without one of them doing any of the things that the other does. Conversely, the professional footballer at work does a great number of things very similar to things done by the amateur footballer who is not working but playing. There are no

ingredient activities common and peculiar to gardening or to working — or to thinking.

Now if someone was under the impression that there did exist some such ingredient activities common and peculiar to gardening or working, he would be forced to allow that they were extremely difficult to isolate. I suggest that part of the difficulty that the experimental psychologists have had in isolating any ingredient acts or states common and peculiar to thinking is just this same difficulty — that of isolating something which is not there to isolate.

Let us just briefly notice a very few of the many different things that we class as thinking. I am thinking if I am going over in my head the courses that I had at last night's banquet. Here there is no problem to solve, no decision to reach. I am thinking, but I am not trying to think anything out. It is more like reverie than like excogitation. On the other hand, if I am doing a piece of multiplication, I am trying to think something out. Yet here there is no room for inventiveness, cleverness or inspiration. There is just a piece of drill, which I can get right or else bungle, and which I can do swiftly and easily or only slowly and laboriously. I am not in any degree puzzled or befogged, and I know all the time exactly what I have got to do and how to do it. There is here no struggling for the right word or phrase, no question of trying to capture the half-formed thought, no place for the bright idea, no room for flair. Compare with this my thinking when trying to translate a Latin poem into an English one. Here there is no question of trying to excogitate an unknown truth. I have a task, but no question, and the solution of my task, if I get it, is not a proposition but an English poem. My problem is how most faithfully and effectively to say something, when I already know, in Latin, what it is that I want to say in English. Here there is plenty of room for inventiveness, inspiration, and flair but little or no room for anything like ratiocination. There is no passage from premisses to conclusion. Suppose, lastly, that yesterday I was set a complicated puzzle, and either thought out the answer or was given it by someone else. Today I go over the solution in my head, moving along the steps which lead to the answer. Here I have no live problem, for I got the answer yesterday. I am not even unsure of the steps. I made sure of them yesterday.



But I am attentively re-tracing an argument, in somewhat the same way as I may go over in my head the courses at yesterday's banquet. Am I thinking or not? Of course, I am thinking, just as I am walking, though not exploring, if I repeat today a walk that I took for the first time yesterday.

We see then that the word 'thinking' covers some activities which are attempts to reach the answers to questions, as well as others which are not; some activities in which there is scope for originality and insight, as well as others where there is not; some activities which incorporate ratiocination, as well as others which do not; some activities, like multiplication and translation, which require special training, as well as others, like reverie, which do not. To look for some common and peculiar ingredients of all thinking is like looking for an ingredient common and peculiar to cats-cradle, hide-and-seek, billiards, Snap and all the other things which we call 'games'.

(4) But now for a point of quite a different sort. People breathe, digest, and grow arthritic however little or much they know about respiration, digestion and arthritis. But people do not play cricket without knowing a lot about cricket — and this is not because they first find themselves playing cricket and then start investigating what they are doing. To play cricket is to do a variety of things all of which one has to learn to do. Cricket is a complex of knacks and techniques, or of drills and skills. It is a truism that a man cannot play cricket who does not know how to play cricket, and what he knows is all that cricketing consists of. There are no hidden ingredients of cricket, though there are all sorts of inevitable and fortuitous concomitants of playing cricket, like panting and perspiring.

Now some, but not all of the things that come under the heading of 'thinking' are in partly the same way complexes of drills and skills. I once had to learn the drills of multiplication and division, and without this training I would not have been even an inefficient multiplier or divider. I should not have multiplied or divided at all. Again, I had to learn how to translate from English into Latin and Greek, and from Latin and Greek into English. This schooling imposed a lot of sheer drills upon me, but it also developed some skills as well.

Consequently the thinking which I do in multiplying and translating is something which I could not conceivably do

without having learned and not forgotten how to do it. I cannot *not* know the knacks, drills and techniques of the computing and translating that I do (though it does not follow that I am as competent to tell other people what they are, as my schoolmasters had to be). Knowing these is knowing what computing and translating consist of. There are no further, concealed ingredients to look for. In some sorts of thinking, like philosophizing and composing, the place of drills, wrinkles and prescribable techniques is much smaller than in computation and translation. To teach a student to philosophize, one cannot do much save philosophize with him. The notion of a welltrained philosopher or poet has something ludicrous in it. But philosophizing and composing are largely without prescribable techniques not because, like panting, perspiring and digesting, they go on so automatically as to be below the level even of being easy; but because to be successful in them is to advance ahead of all the beaten tracks. They require not manuals but practice, stimulation, hard work and flair. There are, patently, lots of kinds of thinking which have something but not everything in common with computing, something else in common with composing, something else in common with philosophizing and so on and so on. The judge may have a very complex problem to solve involving interpretation of law, elucidation of technicalities, unravelling skeins of variegated testimony, and keeping the essentials of the issue steadily before the jury. He may be varyingly good or bad at each of these (and lots of other) more or less dissimilar kinds of thinking. But he is occupied in all of them at once. Similarly the Bridgeplayer's ponderings are, as the Snap-player's are not, an amalgam of highly diverse kinds of considering, at some of which he may be relatively good while he is relatively bad at others. I daresay most of the stretches of thinking which occupy us are mixtures in this general way.

The point is this. To put it for the moment much too bluntly, thinking is an art, like cricket, and not just a natural process, like digesting. Or, to put it less bluntly, the word 'thinking' covers a wide variety of things, some, but not all of which embody, in differing degrees and respects, such things as drills, acquired knacks, techniques and flairs. It is just in so far as they do embody such things that we can describe someone's thinking as careless or careful, strenuous or lazy, rigorous or

loose, efficient or inefficient, wooden or elastic, successful or unsuccessful. Epithets like these belong to the vocabularies of coaches and umpires, and are inapplicable to such natural processes as digesting. We cannot be clever or stupid at digesting, nor yet conservative or independent. (It would be interesting to consider how far epistemologists and psychologists have, unwittingly, yearned to describe thought after the model of digestion.)

Notice that I am not saying that stretches of thinking and games of cricket are not processes. Of course they are. Nor am I saying that thinking and cricketing are unnatural, in any frightening sense of the word. It is quite natural for people to multiply, translate, and theorise, just as it is quite natural for them to play cricket. All that I am saying is, that people, like dogs and lizards can digest without knowing anything about digestion; they can digest whether awake or asleep, infantile or adult, lunatic or sane; but multiplying, translating and theorising, like cricket, have to be learned, and practised; people have to acquire a liking for doing them, and to attend to what they are doing when doing them. To be able to do them is to know what they consist of. Notice, too, that not all the things we class as thinking are subject to the epithets of coaches or umpires. A man in a daydream is thinking, but he is not daydreaming hard, efficiently, rigorously or successfully: nor yet is he daydreaming inefficiently, loosely, carelessly or unsuccessfully. He is not navigating well or badly; he is just drifting.

Notice, lastly, that I am not arguing that there is nothing in thinking that needs to be explored by psychological and physiological methods. The researches of Galton, Henry Head, Freud and Sherrington have led to new knowledge and will lead to more. It is the search for some ingredients or mechanism, whether introspectible or unconscious, common and peculiar to all that goes by the name of 'thinking' which seems to me to be a search for a will o' the wisp. My conclusion is that the experimental investigation of thinking has been, on the whole, unproductive, because the researchers have had confused or erroneous notions of what they were looking for. Their notions of what they were looking for were confused or erroneous partly because they were borrowed from the official philosophical doctrines of the day. They were the heirs of conceptual disorders.

To get the conceptual disorders out of one's system what is needed is not hard experimental work but hard conceptual work.

#### ZUSAMMENFASSUNG

Der Autor wirft zwei Fragen auf: Warum haben die zuverlässigen Untersuchungen der Denkpsychologen (namentlich die der Würzburger Schule) so wenig Einfluss gehabt? Welche Vorstellung haben die Forscher von dem gehatt, was sie untersuchten?

Die funktionellen Beschreibungen und Analysen der von Logikern aufgestellten Theorien über das Denken bildeten nach Ryle den Ausgangspunkt der Untersuchungen der experimentellen Psychologen. Da diese Beschreibungen häufig in der Form introspektiver Mitteilungen wiedergeben worden waren, führten sie die experimentellen Psychologen irre: sie suchten systematische Daten über Elemente, deren Existenz nur auf aprioristischen Gründen festzustellen sind, zu erhalten. Ausserdem wurden von den experimentellen Psychologen Fehler methodologischer Art begangen. Sie behaupteten bestätigt zu haben, dass das Denken eng mit der Aufgabe oder dem Problem verbunden ist. In Wirklichkeit jedoch legten sie ihren Versuchspersonen Probleme vor, von denen sie bereits wussten, dass sie wohl mittels des Denkens gelöst werden konnten, nicht aber ohne das Denken. So nahmen sie also dasjenige, das sie erst beweisen sollten, zum Ausgangspunkt ihrer Untersuchung. Sie gingen ferner von der Annahme aus, dass das Denken eine spezielle Aktivität sei, die in isolierbare Elemente zerlegt werden kann. Sie versuchten also etwas zu isolieren, das seinem Wesen nach nicht zu isolieren ist. Der Autor vertritt die Auffassung, dass das Denken eine „Kunst“ wie das Kricket ist und nicht ein „natürlicher“ Prozess wie die Verdauung. Insofern das Denken eine Kenntnisse erfordert, kann man auch von einem genauen oder ungenauen, einem effizienten oder uneffizienten Denken reden.



Department of Psychology, University, Oxford

SOME REFLECTIONS UPON GILBERT RYLE'S  
CONSIDERATIONS

BY

Prof. G. HUMPHREY

I am somewhat in awe of Professor Ryle's great reputation in philosophy but feel that, with the permission of the Editor, I should like to answer some of the interesting points he has raised.

Taking his first question: "Why has so little emerged from the painstaking investigations of psychologists in the theory of thinking?" Professor Ryle says that he feels guilty because we have started from the philosophers, and experimented in order to find out whether certain philosophical *dicta* were true. There are two considerations here. First of all, experimental investigation has often been made not of a single philosopher's conclusions, but to decide between the conclusions of disagreeing philosophers. To take the example which Professor Ryle gives, namely that of imageless thought. Of the great philosophers, as I have said elsewhere, I think it is fair to say that only Hume "states unequivocally that thinking can be described without residue in terms of images". Kant's polemic against Hume is well known. When people of this stature disagree, surely the psychologist is justified in using his own techniques to find, as far as he is able, which of the two seems to him to be right. The problem in question was, of course, left unsolved by the experimental method, for the Würzburgers decided one way, and the Cornell School, headed by Titchener, the other. So that, in this case at least one has to have recourse to the method of analysis to decide between experiments. I think this should gratify Professor Ryle, and perhaps relieve his conscience (or, perhaps one should call it superego!). Logicians and epistemologists have indeed assured psychologists that "thinking is constituted of ideas or images". But logicians and epistemologists have also assured us that "thinking" is not constituted of ideas or images. Experiments have been done with contradictory results, and to decide between them we have to go back to the method of the philosophers.

Among the experiments to which Professor Ryle objects were the ones which professed to show that Association alone was insufficient to explain thinking, but that an *Aufgabe*, or task must also be postulated. These facts were, he thinks, clear to the experimenters before they started. Why then experiment? Of course, if Association is not sufficient you must either supplement it, modify it, or abolish it altogether as an explanatory principle. The Würzburg experiments which Professor Ryle dislikes were in fact an attempt to patch up the conventional and respectable doctrine of Association by adding to it the doctrine of the *Aufgabe*, much as the later Ptolemaics who found something unexplained added one more epicycle. And just as the "Copernican Revolution" in its physical sense abolished the whole doctrine of cycles and epicycles, so, rightly or wrongly, modern psychology is perhaps tending to abolish Association altogether. This may be seen from the Gestalt attack on it, and also from Freud's general hypothesis that mental processes are based on the Wish.

It seems then probable that the Würzburg attempts to patch the old doctrine will not be finally accepted. But that was hardly the fault of the experimenters. One must remember the times in which these men lived. Here was a conventional doctrine (Association) which was inadequate. Here also was a fact which was patent to them (purpose, task, problem or however one wishes to call it). At least, the existence of a factor describable in some such terms seems patent to us. The Würzburgers tried to combine the two. Being experimentalists, they had to experiment. Nothing else would have satisfied them. Of course the fact that they had to experiment does not validate their experiments.

Apart from what may be called this historical and personal reason, there is a further answer to Ryle's question: "Why experiment on something you know?" For the same question could have been asked of Galileo when he dropped the two balls from the top of the tower. He had, in fact, already confuted Aristotle's theory of falling objects by analytical means (and, I believe, a number of others had done so). One takes it that he performed the experiment to convince first of all himself that it worked and, secondly, to convince the world at large. He would have been a cold fish if in the circumstances he had

not tried the experiment, and I suggest that the Würzburgers did their work on Külpe's ten-year-old notion for somewhat the same reasons. They wanted to try it out, and they wanted to convince their colleagues that it actually was so. It appeared to be so, although most contemporary theory was against them. This seems a perfectly natural thing for an experimentalist to do. Unless people had questioned the obvious, I suppose we should still be thinking that the sun goes round the earth, and that nature abhors a vacuum. If you watch a pump working you can see and hear nature abhorring a vacuum! Surely Professor Ryle has not fallen into Descartes' fallacy of thinking that what is clear and distinct is true!

With the further point that many different activities are classed as thinking, I find myself thoroughly in agreement. Of course it is the province of systematic observation to refine and delimit the current descriptive terms. What the physicist means by energy is not the same as what Aristotle means by it, or modern proprietors of patent medicines. Hence, once again, while granting that the term "Thinking" has been used rather loosely, it is the business of us psychologists either to make it more precise or to abolish it altogether. One may again suspect that the latter course is being followed, and, in fact "Reasoning" was first proposed as a more specific term by American workers, and this is now in process of being superseded by that very ugly word "Problem-Solving".

What puzzles me is that Professor Ryle states that thinking is an art, with the implication that the experimental method is out of place where it is concerned. I tremble for the consequences when I assert that the implication is unjustified. Chemistry is an art also. The Shorter Oxford Dictionary speaks of early writers using the term "as an art only, i.e. *practical* or *applied* chemistry". Thinking is an art as Professor Ryle says. It uses various techniques, and different operations are involved in it. The term is also used to describe apparently different processes. One purpose of scientific investigation is exactly to discover whether they all involve common principles or not; that is to say, to find theoretical rules by which the *practice* of thinking works. N. R. F. Maier has indeed claimed to have found by experiment rules by which, again on experiment, the practice of thinking is improved. May one put it this way; as

practiced by an individual thinker, thinking is an art; but this fact does not itself preclude the experimental investigation of the processes which constitute it.

In general, psychologists ought, I think, to agree with Professor Ryle's conclusion that the experimental investigation of thinking has, up to the present, been on the whole unproductive. Also they should agree with him that psychologists had to start with what they found ready to hand, namely, the conclusions of the philosophers. This is my second main point, and it does not take long in stating. Science starts from reflective thinking. The most systematic reflective thinking has been done by the philosophers, and it is naturally said that philosophy is the "mother of the sciences". I do not think that Professor Ryle need reproach himself with the fact that experimental psychology has begun in the same manner as other experimental sciences, namely, by starting from philosophical thinking. We have in the University of Oxford a distinguished physicist who is still known as the Professor of Experimental Philosophy!

#### SUMMARY

The preceding article asked three questions about the experimental psychology of thinking. (1) Why have the experiments made by psychologists on thinking produced such scanty results? (2) Why have these experiments dealt with the obvious, e.g. that when we think we must first adopt an *Aufgabe* or Task? (3) Why experiment at all on Thinking, which is an art in any case, and is a general term covering many different activities? The answers given are: (1) False philosophical assumptions have been tested by experiment and proved to be false. Negative results necessarily seem nugatory. (2) Obviousness is no final earnest of truth; experiment has shown the falsity of many apparently obvious things. (3) It is the duty of psychologists to find what is common to the different activities classed as thinking, and the fact that thinking is an art does not necessarily preclude its examination by experimental methods.

It is only in the last hundred years that psychology has attempted to break away from philosophy; with other subjects, such as physics and chemistry, the separation is complete. "Thinking" has been later than most branches of the subject in effecting the separation.



*Psychological Institute, University of San Marcos, Lima*

## THE PRECISION OF THE "BLACK THREAD METHOD" AND WEBER'S LAW

BY

Prof. WALTER BLUMENFELD

The "Black thread method", which is used as well in physical and chemical as in biological, psychological and other measurements, may be described with the words of Tuttle and Satterly (8) as follows:

"If a set of experimental measurements, such as those of temperature and length of a metal bar, are found to correspond approximately to a straight-line law, they may be plotted as the x's and y's of a graphic diagram, and their irregularities may be eliminated ('smoothed') by drawing the straight line that appears to come closest to all the points. This is called the *black thread method* because the position of the line is decided by making use of a stretched thread instead of a ruler; the thread and the points can all be seen at the same time, while a ruler would hide half of the points if properly placed. Even better than the black thread is a strip of transparent celluloid... with a fine straight line scratched down the middle of one face". The same authors describe explicitly the most favorable procedure for finding the best position of the 'thread' and for obtaining the equation of the straight line in the form  $y = a + bx$ .

In several other books about the methodology of the natural sciences the same method is mentioned, not always without restraint. In his famous "Treatise of practical Physics" (6), F. Kohlrausch says that such intuitive practices may easily lead astray, especially with regard to the extreme points of a curve. J. F. Guilford (5) would admit the method only "when the points fit the line very closely. Hardly any two observers would extend the line in exactly the same position nor would the same observer do so on two different occasions". G. W. Snedecor (7) considers the method as highly subjective, although he seems to think particularly of cases where one does not know yet if the

correct line will be straight or not. Generally the mathematicians seem to be more sceptical with regard to the black thread method than the physicians.

We shall try to determine the degree to which such distrust is justified. As far as we see, the physicists and psychologists, at least, do not deny completely the value of that method. And as a matter of fact, its application is useful and perhaps sometimes necessary at the beginning of the analysis of experimental data, even though it may be inevitable to revise and correct the obtained result by means of more exact mathematic tools. In such cases the black thread method may be helpful for suggesting a better hypothesis than that of a linear function.

The general procedure for finding the equation of the curve which comes closest to the experimentally given points, is that of the least squares, the knowledge of which we suppose. For our actual purpose we shall only consider the case of the straight line, when the y-observations are likely to have accidental error, the x-observations being strictly correct.

If  $y = a + bx$  is the best straight line through the  $n$  points with the coordinates  $x_1 y_1, x_2 y_2$ , etc, the value of the two constants  $a$  and  $b$  are to be found by means of the following equations:

$$a = \frac{\Sigma(x) \cdot \Sigma(xy) - \Sigma(y) \cdot \Sigma(x^2)}{(\Sigma x)^2 - n \cdot \Sigma(x^2)};$$

$$b = \frac{\Sigma(x) \cdot \Sigma(y) - n \cdot \Sigma(xy)}{(\Sigma x)^2 - n \cdot \Sigma(x^2)}.$$

These equations enable us to determine the exact position of the "regression line", with which the black thread must coincide in order to come closest to all the points. We may, therefore, use the theoretical values of  $a$  and  $b$  as a standard for the comparison with the empirical place of the thread which a subject chooses in view of a series of given points.

The above quoted opinions of some mathematicians and physicists suggest the idea that the validity of the black thread method has been systematically studied. Actually we have not found any evidence of such an investigation in the bibliography we dispose of, and which admittedly is rather incomplete because of the well known situation in most South American libraries. It is however not unlikely that such an examination does not exist at all. For in view of the obvious superiority of the ma-

thematical calculus, the express proof of the inaccuracy of any intuitive method might seem superfluous. More surprising is the apparent absence of psychological investigations with regard to that method, because similar problems have been dealt with as well from a theoretical point of view in the former century, as from the psychotechnical one in the twenties of this century. Later on we shall try to explain the probable lack of such an inquiry, which would justify our present study.

We shall try to examine the precision of the black thread method under definite conditions and to elucidate the psychical processes which occur in its application. More explicitly we shall deal with the following problems:

1. — Which objective accuracy has the subjectively correct straight line when traced between 20 points which cluster in an irregular way around such a line?

2. — Which influence has the mean objective dispersion (deviation) of the given points upon the accuracy of the position of the black thread?

3. — Does the orientation of such a series of scattered points in relation to the subject influence the determination and the accuracy of the tracing of the thread?

4. — Are there great individual differences to be found? And does the personal experience, obtained with general practical research work in the natural sciences, influence the precision of the tracing of the straight line?

5. — Is it possible to establish, on the basis of the results of our experiments, a relation to Weber's law?

6. — Do the psychological processes which occur in the judgments of the subjects, justify such a relation?

Obviously our investigation is partial: one might, for example, ask for the influence of the number of the given points and of the different forms of a regular distribution of them. The straight line being only the simplest case, one might consider other curves likewise. And instead of the effect of general experience in scientific work, one might be interested in that of a specific practice with similar problems.

The empirical and more or less incidental data to which the physicist applies the method of the black thread, such as are given in his measurements of temperature, extensions, time, etc. generally in the positions of the pointers of his instruments,

are not convenient for our purpose. We preferred the construction of a kind of material which has served us already long time ago for different aims (1).

10 cards of white drawing paper were prepared with an original area of  $24 \times 9 \text{ cm}^2$ . In the middle of each we drew a fine straight line (later to be erased), in which 20 points were designed, with a distance of one cm from one to another, leaving at one side of the card 2 cm, and at the other 3 cm free. Perpendicular lines were raised in each point, the length and direction of which to one side or the other of the middle line corresponding to a pre-established plan. The extreme points of these lines were surrounded by black India ink circles, the diameter of which was equal to 3 mm. Afterwards all the mentioned straight lines were erased; the card showed then only 20 small circles irregularly scattered around the original and now invisible central line, to which we shall refer as the "ideal line".

The distribution of the points was determined according to the following norms:

1. — The *algebraic* sum of all their distances from the ideal line was equal to zero, which of course does not imply that the same quantity of points lay at both sides of that line of reference.
2. — The sum of all the *unilateral* distances varied systematically being equal to 20, 30, 40, 50 and 60 mm. We constructed two different cards for each of these distances, which will be referred to as "20 u", and "20 v", "30 u" and "30 v", etc. Our 10 cards were, in that way, graduated with reference to the sum of distances of the points from the ideal line.
3. — The distribution of the points was intended to be completely irregular, in order to avoid as far as possible the appearance of familiar or other "good" figures, such as regular waves, which we had formerly found to produce strong "illusions" (1, p. 376).
4. — The maximum distance of a point from the ideal line was defined as 0,25 times the total unilateral sum. For the card "40 u" as well as for "40 v", for example, that distance would be equal to  $40 \times 0.25 = 10 \text{ mm}$ . The sum of all the remaining unilateral distances would then be  $= 30 \text{ mm}$ . Their size varied of course irregularly.
5. — From the mathematical point of view, the condition N° 1 is not sufficient. If the theoretically "best fitting" straight line



has to coincide with the assumed "ideal" line, forming the x-axis, not only the algebraic sum of the distances of all the points must be equal to zero, but furthermore the sum of their squares must be a minimum. We had indeed to make some slight corrections in our first planned distribution of the points in order to satisfy very approximately this condition. To insist in an absolute accuracy would have been spurious, because we could not draw our points with a greater precision than 0,5 mm. As a matter of fact, the constant "a" in the equations of the calculated straight lines of all our cards is in no case greater than 1,2 mm (determined at the border of them), and the value of the slope (constant "b") is always inferior to .01, which corresponds to an angle of approximately  $0^{\circ}34'$  with the ideal line. The means of the absolute values of "a" and "b" for all our 10 cards are .65 mm and .00514 respectively.

As the length of all our cards was equal to 240 mm, we measured the value of the slope of the thread directly by the distance "t" 240 times greater than "b" and numerically equal to the algebraic difference of the distances from the ideal line, which resulted at the two extremes of our cards. In that way it was possible to avoid the very small values of "b" and to work directly with the measured magnitudes. It was necessary to define the "normal" orientation of the card and the positive sign of the slope. The tangent "b" was considered as positive, when the experimentally resulting straight line (black thread) was turned clockwise with reference to the theoretical one in the normal orientation of the card.

For psychological reasons it seemed appropriate to change the form of our cards in such a manner that the subjects could not imagine the intentional direction of the ideal straight line, which originally had been constructed as objectively parallel to the larger sides of the rectangular cards. Therefore we cut the border of the longer sides of them in an irregular way, as is shown in the Figure 1, which corresponds to our card 60 u, allowing at the same time to appreciate the aspect of the greatest dispersion of the points with which we have worked.

At its small sides two millimetrical rules will be seen, which we used for fixing the position of the black thread with a precision of half a mm. It should be noted that the theoretical zero points do not coincide with those of the scales, nor are these

related to each other. But of course the experimenter knew the theoretically correct values.

Although we have spoken always of a black thread and will do so in the following lines, actually we used a transparent glass

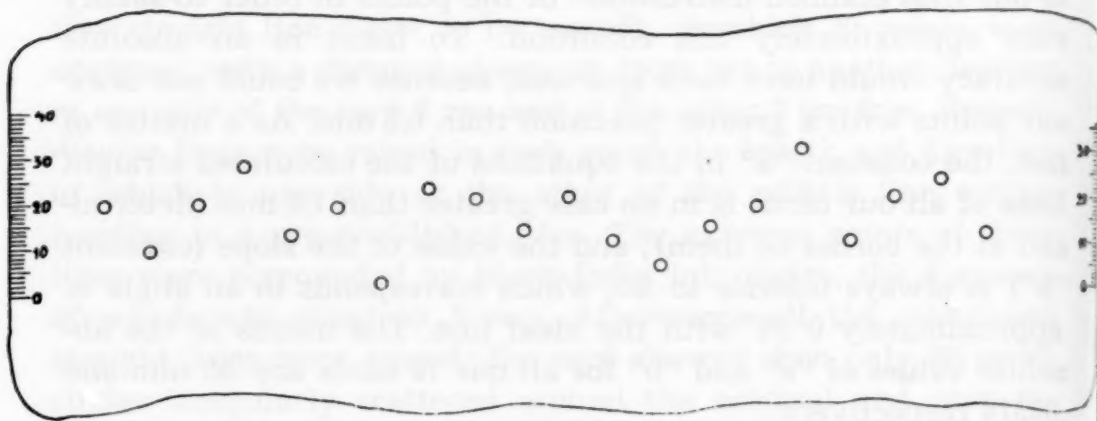


Fig. 1.

plate, 7 cm broad, with a fine straight line scratched not exactly in the middle of the surface and blackened by ink. This instrument was much larger than the cards, allowing the subjects to handle it freely. Yet the experimenter took care that the millimetrical scales at the borders of the cards remained always covered during the activity of the subjects, in order to avoid the influence of the memory and any undesired criteria. The scratched line touched of course always directly the surface of the card, which rested upon a plane drawing board or a glass plate.

The author wishes to express his cordial thanks to his 9 subjects, three of whom were expert scientists, whilst the other six were psychologically well gifted university students.

The experimental procedure was simple. The subjects were informed about their task, namely to find in each of the cards the trend of that straight line around which the visible points seemed to cluster most narrowly, and told that greatest precision and care was expected. Then they were shown one of the cards and the way to handle the thread. They were warned that they should rely on the visual aspect of the points on both sides of the thread without counting them and without trying to estimate the distances from it. It might be as well as not that the thread passed through one or some of the points.

Each card was presented to every one of the subjects 12 times

in two sessions, with an interval of one week, and in a completely irregular order, which changed from one series to the other. The orientation of the cards with reference to the subject varied likewise; they were sometimes turned round, but in such a way, that their large sides lay always approximately parallel to his forehead. 6 series were offered in each orientation. The experimenter noted always with a precision of half a millimeter the places of the lateral scales through which the thread eventually passed.

The computation was, then, easy. We calculated first the means of "a" and "t" for each of our 9 subjects and each card. Afterwards the values related to each card were combined. The "experimental means" which in this way resulted, as well as the "experimental SD" of the 9 individual means, are shown in the Table 1, below the constants of the theoretically correct equations for each card.

TABLE 1.  
Cards

Const.	Value	20u	20v	30u	30v	40u	40v	50u	50v	60u	60v
"a" Theoretical	— .3	+ .7	— .2	— .3	+ 1.0	— 1.0	— .4	+ .8	— 1.2	+ .6	
exp. mean	— .2	+ 1.2	+ .2	+ .6	+ .4	— .5	+ .5	+ 1.5	— .9	+ .9	
exp. SD.	.34	.53	1.0	1.08	.65	1.47	.73	1.94	2.04	1.58	
		.44		1.04		1.06		1.33		1.81	
"t" Theoretical	+ .23	— 1.40	+ .16	+ .31	— 2.18	+ 1.87	+ .58	— 1.89	+ 2.30	— 1.41	
exp. mean	+ .19	— 2.54	— 1.63	— 1.60	— .85	+ .90	— 1.31	— 2.78	+ 1.41	— 1.73	
exp. SD.	.68	1.04	1.36	2.21	1.36	2.87	1.58	2.67	4.46	2.91	
		.86		1.78		2.12		2.13		3.69	

Theoretical and empirical values of the two constants of the straight lines. The means refer to 9 subjects, each of whom determined the thread 12 times for each card. The values of SD correspond to the variation of the individual means.

As one sees immediately, the differences between the theoretical values of "a" and "t" and the experimental means are small, the greatest deviation with reference to "a" being equal to .9 mm, and with reference to "t" equal to 1.9 mm. The experimental values of SD are not considerable either. The maximum value of  $SD_a$  is 2.04 and that of  $SD_t = 4.46$ .

With regard to the individual means, which do not appear in the table, it must be said that there exist rather great differences if one considers the numerical relation of the absolute values of both constants. But the differences themselves are fairly small.

Amongst 90 cases (10 cards for each of 9 subjects), the deviation of "a" from the theoretical value is only 32 times greater than 1 mm, and only 10 times greater than 2 mm, the maximum being equal to 4.2 mm.

In order to estimate the magnitude of the constant "t", it should be considered that for obtaining the tangent of the angle, the values of table 1 must be divided by 240. All the angles are, therefore, very small.  $t = 4.2$  mm corresponds to an angle of  $1^\circ$ . Actually the distances "t" are only 10 times (among 90) greater than 4 mm, and the maximum which has happened is equal to 8.77 mm, the slope being then  $2^\circ 5'$ . The table shows that the mean values come very close to the theoretical ones.

It is strange that the differences between the theoretical values of "a" and "t" and the experimental means have no clear relation to the degree of dispersion of the points of our cards. To be sure, the most accurate position of the thread, as measured by "a" and "t", occurs with reference to our card 20 u, but the least precision corresponds to the cards 30 v and 50 u, instead of 60 u and 60 v. If we compare the dispersion of the points in all our cards with the accuracy of the empirically found constants, the Spearman-coefficient of rank correlation is low, namely  $+ .048$  for "a" and  $- .127$  for "t".

The variation of SD agrees better with the dispersion. Especially if one forms the mean values of the SD values, which belong to two cards of the same degree of dispersion, as has been done in the 4th and 8th line of the table 1, one notes an approximately linear progression, which might be expressed by the two equations:

$$SD_a = - .078 + .0304 x;$$

$$SD_t = - .298 + .0603 x.$$

The result of our investigation seems to be that the method of the black thread does not deserve so much distrust, as the sceptical judgments might suggest which we have quoted before. This holds at least true under the particular conditions of our experimentation and of its evaluation. These conditions are in part favorable for obtaining good results: We have worked with patterns of 20 points; with a smaller number the outcome might have been worse. Furthermore, we have considered the means of 12 experiments for each subject. Later on we shall see, to be sure, that this number might have been reduced to the half



without affecting substantially the accuracy; anyhow, this accuracy keeps a statistical character. Finally, we have eliminated or at least diminished the influence of such arrangements of the points as would possibly cause "illusions" and which in a real investigation might occur. But this last argument would likewise apply to the mathematical treatment of the data, falsifying the results. On the other hand, we have introduced also some rather unfavorable conditions, because we abstained from using ruled paper for the cards and because at least the cards 50 u and 50 v, as well as 60 u and 60 v show a very high degree of dispersion.

As has already been told, the individual differences between our subjects are *relatively* great. If one considers only the straight lines chosen by our 5 "best" subjects, the differences between the mean values of their constants and those of the theoretical straight lines become quite insignificant. The greatest deviation of the constant "a" is then equal to .6 mm, and that of the constant "t" equal to 1.03 mm, which corresponds to an angle of  $0^{\circ} 10,6'$ .

One might expect that amongst these 5 best subjects all our 3 scientists are to be found. Actually this is not the case. One of them presents even the greatest variation of all our subjects. And the greatest accuracy corresponds to one of our young university students. The general scientific practice seems, therefore, not to be so decisive for the precision in tasks akin to ours, as some individual gift. So far as the present writer could observe, all his subjects have worked very carefully, without haste and to the best of their ability.

As has been mentioned, the cards were presented to our subjects in two opposite orientations, the change of which they could not perceive. When comparing the means of the constants "a" and "t" for our five most sensible persons, it becomes clear that the orientation has no substantial influence. The greatest difference which occurred with regard to "a" is equal to .7 mm, and with regard to "t" equal to 1.1 mm. Hence the inversion of the cards does not essentially change the appearance of the series and the phenomenal trend of the points. It would have been possible to reduce the number of our individual experiments to 6 for each subject, without impairing the accuracy of the means.

So far we have been exclusively interested in these values.

But their fluctuations are likewise important as well for the scientist who applies the method, as for the theoretical psychologist. We have already seen that the degree of conformity of our subjects increases according to a linear function with the dispersion of the points of our cards. But as Guilford has said (5), the position of the thread changes more or less when the same subject repeats the task in the face of the same card. We shall therefore now consider the magnitude of the individual variances with reference to the degree of dispersion of the points. As usual, we determine that magnitude by the respective standard deviations  $SD_a$  and  $SD_t$ .

TABLE 2.  
Cards

Constant	20u	20v	30u	30v	40u	40v	50u	50v	60u	60v
"a"	.82	.90	1.17	1.33	1.62	1.60	2.06	1.96	2.24	2.66
	.86		1.25		1.61		2.01		2.45	
"t"	1.38	1.54	2.16	2.80	2.97	3.49	3.69	3.86	4.36	5.16
	1.46		2.48		3.23		3.77		4.76	

Mean values of the standard deviations which show the different subjects with respect to the two constants. In the second and fourth rows the values for the cards of the same degree of dispersion have been combined.

In the table 2 we have combined the arithmetic means of these values of all our subjects. They should be distinguished from the SD values which exist in the table 1, for there we dealt with the differences of the means of the two constants in all our subjects. But the values in the table 2 have been found in another way. We have first calculated the standard deviations of the constants "a" and "t" which occur in each subject with regard to each card. Afterwards we have computed the means of these values for all our subjects. Only these means are shown in the table.

As a matter of fact, the *relative* individual differences are considerable. The SD of one subject is in an extreme case 7 times that of another with regard to the same card. The 5 persons whose mean values of "a" and "t" came closest to the constants of the theoretically correct line, are at the same time those whose constants fluctuate least. But with the exception of one subject, the *absolute* magnitudes of SD, are small, and so are the numbers which appear in our table 2. It is somewhat surprising that these values differ so much in some pairs of cards, which have the same objective degree of dispersion. We

have combined them in the respective even lines of the table, where their means are to be seen.

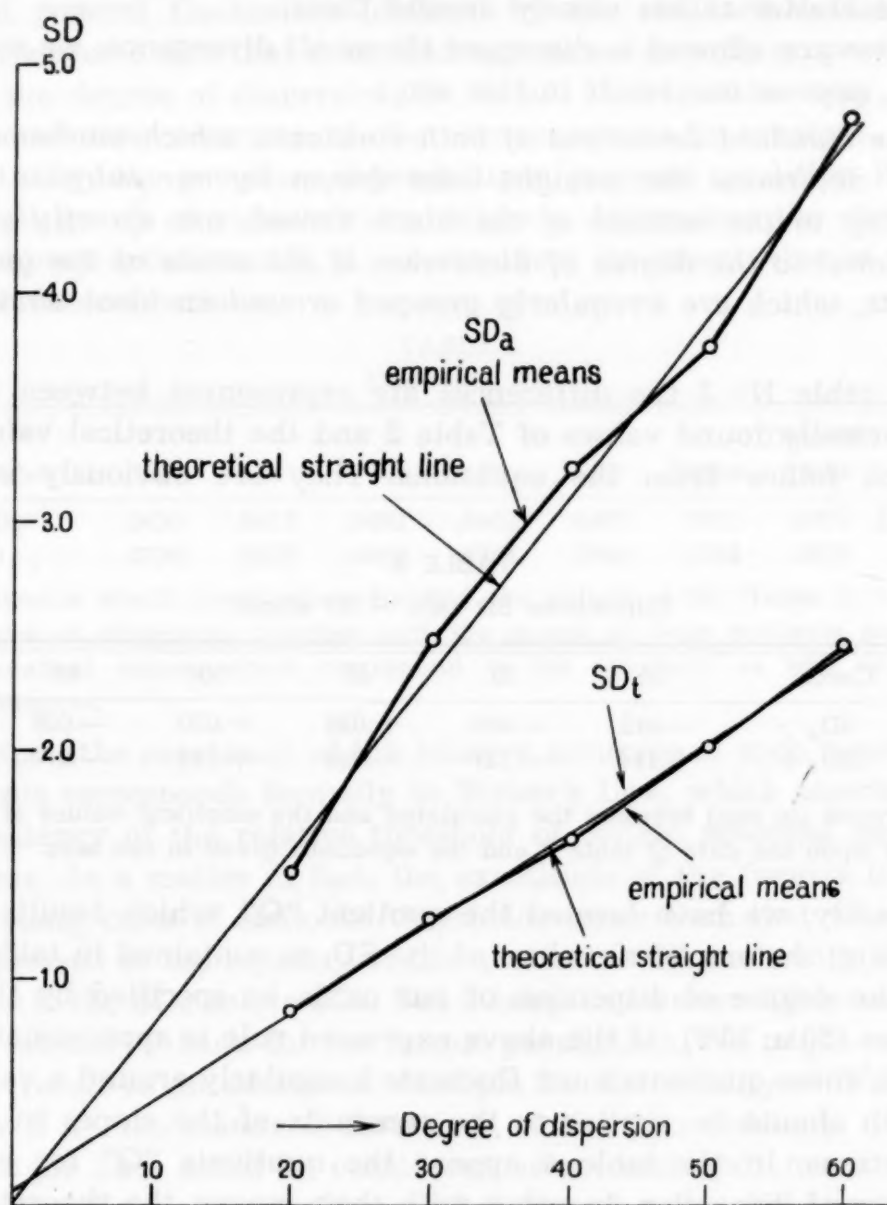


Fig. 2.

The general trend of these values is very clear; they increase in function of the degree of dispersion. In figure 2 the two lines have been represented together with the best fitting straight lines, the equations of which have been derived in the usual way. They are as follows with all the measurements expressed in mm:

$$SD_a = .060 + .0394 x;$$

$$SD_t = .016 + .0779 x.$$

As is easy to be seen, the theoretical lines cross the y-axis in points very near to the origin, and the empirical points of the curve cluster rather closely around them.

If we are allowed to disregard the small divergence, we may, then, express our result in this way:

*The standard deviations of both constants, which mathematically determine the straight lines drawn by our subjects according to the method of the black thread, are directly proportional to the degree of dispersion of the series of the given points, which are irregularly grouped around an ideal straight line.*

In table N° 3 the differences are represented between the empirically found values of Table 2 and the theoretical values, which follow from the equations. They are obviously very small.

TABLE 3.  
Differences SD calc. — SD empir.

Cards:	20	30	40	50	60
SD <sub>a</sub> :	— .012	— .008	+ .026	+ .020	— .026
SD <sub>t</sub> :	+ .114	— .127	— .098	+ .141	— .070

Difference (in mm) between the calculated and the empirical values of SD, based upon the data of table 2 and the equations given in the text.

Finally, we have formed the quotient "Q", which results by dividing the empirical values of the SD, as contained in table 2, by the degree of dispersion of our cards, as specified by their names (20 u, 30 v). If the above expressed rule is approximately valid, these quotients must fluctuate irregularly around a value, which should be similar to the constants of the slopes in our equations. In the table 4 appear the quotients "Q" for each degree of dispersion, together with their means, the theoretical value of the slope and the difference.

The difference between the mean of Q and the theoretical value is very small indeed. We may, then, consider the linear dependency between the SD and the degree of dispersion as well established, under the conditions of our experiment.

This result agrees only in part with the warnings of the physicists and psychologists, who assert that the method of the black thread can be relied upon only when the empirical points



cluster narrowly around the supposed straight line. Now, this holds true with regard to the absolute amount of the individual and general fluctuation (defined by the standard deviation); but we have seen that these fluctuations are directly proportional to the degree of dispersion, or, to put it in another way, that *the relative accuracy of the measurements is constant*. Admittedly, the precision depends upon personal qualities. But a small group of conscientious subjects determines the straight line very satisfactorily, if one considers the mean value of some few determinations.

TABLE 4.  
Quotients "Q"

Cards:	20	30	40	50	60	Mean	Theoret. value	Diff.
$SD_a$ :	.0430	.0417	.0403	.0402	.0407	.0412	.0394	.0018
$SD_t$ :	.0730	.0827	.0808	.0754	.0793	.0782	.0779	.0003

Quotients which result when dividing the values of SD (Table 2) by the degree of dispersion, together with the means of these quotients and the theoretical values which correspond to the equations as well as their differences.

Now, the constancy of the relative accuracy of such measurements corresponds formally to Weber's Law, which asserts the constancy of the relative threshold of certain sensorial experiences. As a matter of fact, the exactitude of the famous law is in many cases of sensorial thresholds lesser than that which is expressed in the equality of the quotients "Q", both with reference to the  $SD_a$  and to  $SD_t$ . But it keeps to be seen if there exists an intrinsical basis for the formal parallelism. We might accept our result as an additional example for the validity of Weber's law if 1) the standard deviations of the two constants of the straight lines could be considered as a (combined) measure of the threshold of its position; and if 2) such a threshold could be considered as something like a "stimulus" or an attribute of a stimulus.

Let us consider the second point first. Weber's law has been found valid, within certain limits, for the intensities of the sensorial stimuli and, with certain restrictions and exceptions, for many qualities and extensions in space and time. One might discuss if all these "objects" should be considered as "stimuli" in the classical sense of the word. Boring, Langfeld and Weld (2)

define this concept as "any change in external energy that gives rise to such an excitation of the nervous system as arises a response" (not only those "which affect the tissues directly") or, more specifically, such changes in external energy as "activate a sense organ and its receptors". Differential thresholds would, then, imply a just perceptible difference of the magnitude of such changes.

It is not easy to apply these definitions to all the cases which are usually considered as belonging to the sphere of Weber's law. One might question, for example, at least from the standpoint of the "classical" psychology, that the duration of a time-interval limited by two clicks, is a stimulus and that the difference between two different time-intervals involves a change of external energy. Much the same could be said with regard to the lengths of two distances, especially if not drawn in the paper, but only demarcated by two small strokes or dots. In order to uphold the definition, one would have to resort to the whole physiological state of the respective sector of the central nervous system and to hypothetical changes of energy in it.

To be sure, the behaviorists apply the term "stimulus" in a different way. J. F. Dashiell considers the "simple stimulus", which is a form of energy leading to a simple response, as a convenient abstraction only, because as a rule there would exist a "multiple stimulus leading to a multiple response" (4, p. 41). And so complex things as a word, a phrase, a sentence "may become an effective stimulus". A waving white glove of a policeman or the traffic signal is a stimulus for the chauffeur, as well as the appearance of the mother is a stimulus for the child. In such cases it is evident, that the concept "stimulus" has at best a very loose reference to the energy changes, and a very intimate one to the symbolic value of an object or a situation. We wonder if anybody would expect that Weber's law might be applied to such "stimuli".

We shall not try to analyze and determine more exactly the use of the problematic concept of stimulus, especially because the Gestalt Theory has greatly changed the aspect of all the questions related to "sensation" and "stimulus". In our particular case, we deal with visual figures which consist of a series of irregularly dispersed points. Obviously the degree of irregularity is a quality which does not belong to any of the points or to

a definite part of them, but to the *whole* series as such, comparable to the symmetry of a façade or to the character of "major" and "minor" in musical themes and accords. When looking at one of our cards, one has immediately the impression of certain "intranquility", and it is remarkable that this quality has *degrees*. Two series of completely different construction may have the same degree of such intranquility, irregularity or "agitation", although no point of one series must have exactly the same distance from the "ideal line"; and on the other hand, we have shown, long time ago, that one may change fundamentally the degree of phenomenal irregularity by changing only the *order* of the points of a series, whilst all the distances of the "ideal line" remain exactly the same (1, p. 377).

It would be very risky to assert that to a series of greater irregularity of 20 points, there must correspond a greater amount of external energy which acts upon the nervous system. The hypothesis seems more plausible that the special character of the global excitation of the visual sector in the brain is related to the phenomenal degree of dispersion, which is shown by the points of our series, in agreement with Wertheimer-Koehler's principle of isomorphism. These excitations would then form the basis for the comparison of two series.

In our former investigation, we had determined the value of the relative differential threshold of such irregular configurations of points as have been used in our actual experiments. Weber's constant was found to vary between  $1/12$  and  $1/15$  approximately, the greater value corresponding to the smaller dispersion of the points. It is also interesting that those subjects likewise showed considerable differences in their sensibility.

It might seem that our actual experiments had apparently nothing to do with the determination of differential thresholds, but with that of the position of the black thread and its precision. Such an opinion would, however, be wrong. For our own observations as well as those of our subjects show, that the procedure of finding the correct straight line consists in a continual sequence of comparisons. The subject who turns and moves the thread to one side or the other, balances incessantly the two parts of the series of points which lie on both sides of the thread, until he feels satisfied by the phenomenal equilibrium between them. To be sure, the construction of the cards excludes any

symmetry or equality. One greater distance at one side of the ideal line may be compensated by two or more at the other side, and in another region of the card. The phenomenal equilibrium is, therefore, not so simple as that which is required when dividing a line in two equal parts. It is more like the mechanical one of a lever of the second kind, where many forces act on both sides of the fulcrum, or perhaps also like the balance which the classic painters demand with regard to the shapes, colours or degrees of clearness in different parts of their works.

The process of comparison is essential for the judgment of our subjects, and the accuracy of their measurement depends upon the specific sensibility which they have for the equality of the two "momenta" at both sides of the thread. This is exactly the general problem which we are concerned with when determining a differential threshold. The validity of Weber's law, as results in our experiments, would then be only one confirmation more among many other cases. But it is a rather particular case, which we are dealing with; for at least two reasons:

- 1) It implies an exact numerical determination of the threshold of "gestalten", which are much more complex than those, with which most investigations, since the first attempt of K. Buehler (3), were concerned;

- 2) The precision of the position of a straight line is defined by an equation with two *constants*. Now we hope to have shown that, under the conditions of our experiments, *Weber's law holds true with respect to both of them*. We might put it as follows: The degree of fluctuation of both constants of the equation of a straight line, as measured by the SD in our experiments, is directly proportional to the degree of dispersion of the points, which cluster around the objectively correct straight line.

#### SUMMARY

In order to determine the precision of the "black thread method", 9 subjects were examined with a set of 10 especially constructed cards, which were systematically graduated with regard to the dispersion of 20 points irregularly scattered around an invisible straight line. The equations of the best fitting lines being known, the mean empirical constants of the lines which resulted in 12 experiments for each subject and card could be



compared with them. The principal conclusions are the following ones:

1. — The general means of both constants of the empirical straight lines differ very little from the theoretically correct ones.
2. — The individual differences, though *relatively* notable, are on the whole not very considerable.
3. — The accuracy of the work done by expert scientists is not superior to that of university students.
4. — The absolute individual differences increase approximately in linear progression with the degree of dispersion of the points of our cards.
5. — The presentation of the cards in two opposite orientations does not influence sensibly the results.
6. — The means of the SD of both constants in all our subjects increase in an astonishingly exact proportion with the degree of the dispersion of the points.
7. — The psychical processes during the operations of the subjects consist of constant comparisons of both parts of each card which the black thread divides. It seems, therefore, justifiable to extend the validity of Weber's law to configurations (Gestalten) like these, in the sense, that both constants of the equations of a straight line show values of SD (thresholds) which increase in direct proportion to the degree of dispersion of the points.

#### ZUSAMMENFASSUNG

Um die Genauigkeit der „Methode des schwarzen Fadens“ zu prüfen, wurden 10 Karten mit unregelmässig um eine „ideale“ Gerade herum verteilten Punkten konstruiert, deren Streuung systematisch variierte. 9 Vpn. bestimmten je 12mal die Lage einer auf einer Glasplatte eingeritzten Geraden, die subjektiv am besten jeder einzelnen „idealen“ Geraden entsprach.

Fast alle Vpn trafen im Mittel van 6 Einstellungen mit ausreichender Genauigkeit die Gerade, welche durch die beiden Konstanten der Gleichung  $y = ax - b$  definiert ist. Wechselnde Orientierung der Karten zur Vp hatte keinen Einfluss. Immerhin zeigten sich erhebliche individuelle Abweichungen zwischen den nach Alter und wissenschaftlicher Erfahrung sehr verschiedenen Personen.

Zur Bestimmung der Streuung der beiden Konstanten  $a$  und

b dienten die SD — Werte, welche als Mass der Unterschiedsschwelle betrachtet werden können. Die entsprechenden Mittelwerte aller individuellen Bestimmungen variieren, und zwar so, dass die Schwellen jeder der beiden Konstanten für sich genommen direkt und linear vom Grade der Streuung der Punktkonstellationen abhängen. Die geraden Linien, die mathematisch diesen mittleren SD — Werten entsprechen, gehen fast genau durch den Ursprung des Koordinatensystems hindurch. Danach würde das Webersche Gesetz der Konstanz der relativen Unterschiedsschwelle mit grosser Genauigkeit gelten, sofern es zulässig ist, die vorliegende Untersuchung theoretisch in Beziehung zu den klassischen Arbeiten über die Psychophysik sehr viel einfacherer Prozesse zu bringen.

Die Analyse der bei unseren Experimenten auftretenden Prozesse rechtfertigt diese Hypothese. Denn die unregelmässig verteilten Punktgruppen haben die Struktur gerader Linien von variabler „Unruhe“ oder „Bewegtheit“. Physiologisch dürften ihnen globale Erregungen im Zentralnervensystem entsprechen, deren durch die jeweilige Lage des „schwarzen Fadens“ getrennte „Teile“ hinsichtlich des Grades ihrer „Unruhe“ verglichen werden.

Demnach ergeben sich folgende Sätze:

1) Die „Methode des schwarzen Fadens“ ist hinreichend genau, wenn man die Mittelwerte von 6 Bestimmungen verschiedener Personen zugrundelegt.

2) Der Grad der Streuung (SD) jeder einzelnen der beiden Konstanten, welche die Gleichung der Geraden bestimmen, ist direkt proportional der objektiven mittleren Streuung der Punkte, die sich um die theoretisch richtige Gerade scharen. Demnach gilt das Webersche Gesetz innerhalb der Grenzen dieser Untersuchung.

#### REFERENCES

1. Blumenfeld, W., Urteil und Beurteilung. Leipzig 1931.
2. Boring, E. G., H. S. Langfeld and H. P. Weld, Foundations of Psychology. New York 1928.
3. Bühler, K., Die Gestaltwahrnehmungen I. 1913.
4. Dashiell, J. F., Foundations of objective Psychology. New York 1928.
5. Guilford, J. F., Psychometric Methods. New York 1936.
6. Kohlrausch, F., Lehrbuch der praktischen Physik. 12. Aufl. Leipzig 1914.
7. Snedecor, G. W., Statistical Methods. Ames, Iowa 1937.
8. Tuttle, L. and J. Satterly, The Theory of measurement. London 1925.

Department of Zoology, University, Liverpool

## THE ORIGIN OF LANGUAGE \*)

BY

Prof. R. J. PUMPHREY, F. R. S.

The origin of language is a subject of general interest and has long been one. The tower of Babel has its parallels in other cultures. In most folk-lore an understanding of the speech of animals is one of the attributes of sage and hero. Theories of the origin of language have interested persons of inquiring mind in this country at least since the time of Locke whose *Essay* is the first attempt, at least in English, to investigate the mutual relations of language, of understanding and of behaviour. The whole subject is so intricately tied up with the meaning of meaning and the basis of logic that its implications can hardly be avoided in any intellectual study from metaphysics to mathematics.

But, since any theory of language is necessarily a guess and so many brilliant minds have been guessing for so long, it might appear that the best guesses had already been made, and that yet another was entirely redundant, I want to suggest to you, on the contrary, that there is fresh evidence, some of it fairly recent, which has not yet been adequately considered in its relation to the subject of this lecture.

There are two schools of thought about the origins of human speech. All are agreed that human speech differs in material particulars from the speech of other animals. There are those who, like Darwin, believe in a gradual evolution, but there have been others who have believed that speech is specifically a human attribute, a function *de novo*, different in kind from anything of which other animals are capable. Did language evolve or was it invented? And in either event what can be inferred about its history? Evidently before the conflict of opinion can be determined it is necessary to define, or at any rate, to consider how far definition is at present possible of the

---

\*) An inaugural lecture by the Professor of Zoology, The University of Liverpool, 1951.

ways in which human speech is fundamentally different from animal speech, if indeed any other animal can be said to speak at all.

It will be admitted that speech is but one of a number of possible means by which one individual may communicate information to others; and on communication between non-human animals, recent work has compelled a radical revision of our ideas. It was always obvious that neither social life (which is quite common in the animal kingdom) nor sexual life (which is well-nigh universal) would be possible without some communicative faculty, but man, in his vanity, has been apt to consider that it must of necessity be vague and indefinite except, of course, in his own case. Now von Frisch has demonstrated beyond cavil that the returning worker-bee delivers to her colleagues in the hive a message as concise and explicit as a naval signal. It is as if she announced: "At 750 yards on bearing  $186 \pm 1^\circ$  a patch of clover in full nectar...." The message is precise and it is articulate, that is to say, symbolic representations of four things with a common reference, namely the kind and quality of the source of nectar or pollen and the two coordinates which fix its position, are linked in one message.

Now it is noteworthy that the representation of things by symbols and the linking of symbols together to make a coherent message have been considered to be diagnostic of human speech and writing. Of course information *can* only be communicated symbolically so that the use of symbols must necessarily be coextensive in the Animal Kingdom with the faculty of communication, but it could be argued that a particular form or degree of coherence of symbols was characteristically and exclusively human.

It is worth-while, therefore, to consider the language of the bee in a little more detail. Neglecting some exceptional cases and limiting conditions, what happens is as follows. A returning worker who has found a new and sufficiently rich source of nectar enters the hive. Wagging her abdomen, presumably with the purpose of distributing her scent and exciting her colleagues, she traces out in pitch darkness a figure-of-eight path on the vertical wall of a comb. In this progress her speed is an inverse function of the distance of the source of nectar from the hive, and the line bisecting the figure-of-eight makes with the vertical



an angle proportional in magnitude (though, of course, in a different plane) to the angle between the azimuth of the sun and that of the spoil as observed from the hive. The other two items of information, the kind and quality of the nectar, are presumed to be represented by the kind and quality of the odour clinging to the messenger. Such bees in the hive as are interested follow the messenger in her progress nuzzling her flanks, and it has been shown that these bees act on the message accurately if they act at all.

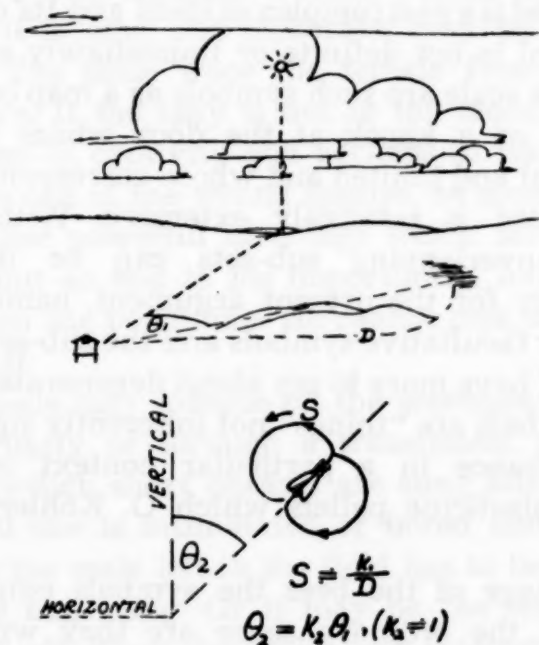


Fig. 1. Illustrates the bee conventions and also two complicating factors.

(1) The proportionality constant,  $K_2$ , is of order unity but varies with the height of the sun. This is probably correlated with von Frisch's subsequent discovery that the bee takes its bearing not from the sun directly but from the polarization of the blue sky. Hence the system works when the sun is overcast (as shown).

(2) If the return flight involves a detour the *actual* flying distance and the *true* bearing are indicated. Hence the correct interpretation of the message involves its integration with topographical memory.

A fact of some theoretical importance (not indicable in the diagram) is that the message is informative but not imperative. It is as if the Ministry of Labour were to let it be known in the Rhondda that there was coal to be hewn in Nottingham. Some miners might act on it, especially if their own pit was closed, but many would not. So bees accustomed to foraging near at hand generally ignore messages about nectar at a great distance. Bees accustomed to collect from catmint may ignore a clover message unless the catmint source is running low, and so on.

Now a symbol may be adequately, if clumsily, defined as something which represents, for the purpose of conveying information another thing with which it is not identical but to which it (generally) corresponds in some particular. "Thing" here must be taken to refer not only to any member of the set of concrete and denumerable objects but to any action, process, relation, experience, in fact to anything.

Symbols in human experience vary enormously. At one end of the scale are symbols such as the crown and the cross, where the thing represented is a vast complex of ideas and its correspondence with the symbol is not definite or immediately evident. At the other end of the scale are such symbols as a map or a photograph or a footprint or a knock at the door whose significance is relatively trivial and limited and whose correspondence with the thing represented is relatively extensive. Within the set of symbols two overlapping sub-sets can be delimited with sufficient clarity for the present argument, namely the sub-set of occasional or facultative symbols and the sub-set of degenerate symbols. I shall have more to say about degenerate symbols later; facultative symbols are "things" not inherently significant which acquire significance in a particular context or *Gestalt*, for example, the plasticine pellets which O. Köhler's birds learnt to count.

In the language of the bees the symbols employed are not sounds, nor in the ordinary sense are they writing. In their employment they correspond rather to a kind of animated cartography in which smells and tactual stimuli replace the little explanatory pictures and scales of distance of the old map-makers. Such a system might at first seem incredibly elaborate and sophisticated for a mere insect, but it is salutary to remember that the civilization of the honey-bee is immeasurably more ancient than our own and that bees were going about their business in an organized polity at a time when our ancestors looked and probably behaved like rather unintelligent rats.

Let us turn now to the work of Albrecht Faber on grasshoppers of the genus *Chorthippus*. Faber was not interested in the theory of communication. He was an ecologist who hoped by listening to the chorus of different species of grasshoppers, to estimate their population density without having to catch and count them. This led him to investigate when and how and why they sing,

the first serious study of their language, though sketchy observations had been made by Yersin and others previously. I summarize his results very briefly, together with some additional observations of my own.

The songs of each species are recognisably different. It is generally only the males who sing, and within each species there are several types of song. Firstly there is the ordinary song which seems to signify that the singer is disengaged and ready for anything. Secondly there is the serenade, addressed by a single male to a marriageable female, which, at least in those species where it is long and elaborate, may be interrupted in various ways. The male woos the female from a station immediately astern. If the lady is not in the mood, or if, as may certainly happen in captivity and perhaps in the field, he pays his addresses to one of the wrong species, he soon gets a kick in the face from her powerful hind legs which sends him flying. This may not put an end to his importunity, but he must now begin again from the beginning. He cannot pick up the serenade where he left off.

Or another male may intrude on the serenade, in which case it changes abruptly or through a transitional passage to the rival's duet in which short phrases are sung alternately by the two males until one is intimidated or bored and takes himself off. Once more the male left in the field has to begin his wooing again from the beginning. Or it may be the female who gets bored and hops off. Then the males lose interest in each other and each reverts to the ordinary song. If ultimately the serenade is allowed to reach its climactic end, it is followed by a kind of shriek of triumph, the *Paarungslaut* which leads immediately to mating; and this process in turn may be accompanied by yet a fifth type of song. There may also be characteristic cries associated with getting ready to jump or elicited by disturbance by an external agent.

These recognisable sound-patterns range in length and complexity from a single interjection or expletive to a piece of music with elaborate scoring. Here we clearly have a language of symbolic sounds but the symbols in this case represent emotional states of the singer and evoke corresponding emotional responses in the auditor. It differs in content from the language of the bees (or, at least, from that part of it which von Frisch has made

us aware of) as much as the "Magic Flute" differs from a news bulletin, but it is evidently a language. Whether it is to be regarded as an articulate language depends on definition. There is a coherent and meaningful sequence of songs. The songs themselves are musically articulate, they convey information and, even to our ears, they are expressive.

The ordinary song, which is simply an expression of readiness, is as monotonous as the dialling tone of a telephone and frequently not unlike it in tempo. The challenge and counter-challenge of the rival's duet is unmistakable. The serenade in those species in which it is well-developed begins quietly with a sound like a watch ticking very slowly and develops crescendo and accelerando on a surge of mounting passion till, as the climax approaches, not only the legs but the whole body of the insect is blurred with activity. And the *Paarungslaut* sounds triumphant.

Now there is evidently an important difference between the language of the bee and that of the grasshopper. The former is a social phenomenon, businesslike and highly specialized in its application to a particular aspect of social life, the latter seems to be entirely or almost entirely related to sexual activity, which is common ground with nearly all animals. The meaning of the bee's language is so far from being obvious that von Frisch's decipherment of it is an achievement which ranks with Champollion's decipherment of the hieroglyphics on the Rosetta stone. The meaning of the grasshopper's language is, I suggest, self-evident.

In making this suggestion I am aware that I lay myself open to an accusation of having indulged in the pathetic fallacy,—of being, in the modern jargon, an animist. If animist meant simply non-mechanist I should take the insult as a compliment, but unfortunately it has acquired a more sinister aura. So I must try to be precise. I am suggesting that a naïve human observer could guess the significance of the grasshopper's songs by listening to them, because it would appear to him "natural" that the ordinary song, where the male is simply a sound beacon for disengaged females, should be monotonous; it would appear to him natural that the exchanges between rivals should be sharp and staccato; it would appear to him natural that the gradual increase in tempo and loudness of the serenade should herald



an emotional climax. And, if this is true, it implies that there is a symbolism common to man and the grasshopper in the use of sound to express and evoke emotion.

Of course it would be absurd to generalize when only one genus of sound-producing insect has been examined with any care. And there is the complication that in many animals auditory symbolism associated with sex is partly or wholly replaced by visual symbolism as in spiders and birds, or osmic or tactual symbolism as in many insects and mammals; and no attempt at collation and correlation has yet been made. Animal semantics is in its infancy if indeed it has yet been born. But if it were found to be true, I do not myself feel that it would be particularly astonishing, for the similarities in the neural organization of men and insects are more remarkable than the differences. And the emotive effects of sound on the human organism are evidently deep-rooted and of long standing. There is ample evidence of this though I have no time to go into it here.

I have taken bees and grasshoppers as examples partly because the facts are perhaps less widely known than facts about birds or mammals and the evidence less clouded by anecdote and tradition, partly also in order to avoid connecting the communicative faculty with intellect. No one has yet defined intellect to the satisfaction of anyone else, but we are all agreed that we have a great deal of this commodity and that insects have practically none. So the ability to communicate information symbolically can not be regarded as a result of intellectual ability.

What is it then which characterises human speech? For it is clear that man cannot claim a monopoly of messages which are precise, articulate, expressive and perfectly adequate to the ends which they serve.

Let us approach the question from a different point. It is often said that speech—human speech—has two contrasting functions, the emotive and the informative, the "Magic Flute" and the news bulletin over again. The terms are not happily chosen because the emotive function is also manifestly informative, and it may often be as important to be able to appreciate whether a voice is charged with hostility or affection or deceit, as to distinguish the words spoken, particularly if they happen to be

spoken in a language one does not understand. I shall hereafter use the term *information* to denote the total of information in a spoken message and the term *intelligence* to denote that part of the information which can be taken down in writing and used in evidence.

The distinction is of considerable importance to telephone companies and its investigation has led to results which are both interesting in themselves and relevant here. It is very difficult to design a telephone which will transmit all the information of human speech and it has long been known that for many purposes it is quite unnecessary. But is it only recently that a serious attempt has been made to see how bad a telephone can be without loss of intelligibility. To explain the results I must refer to the diagram which is a way of picturing the performance of the average human ear

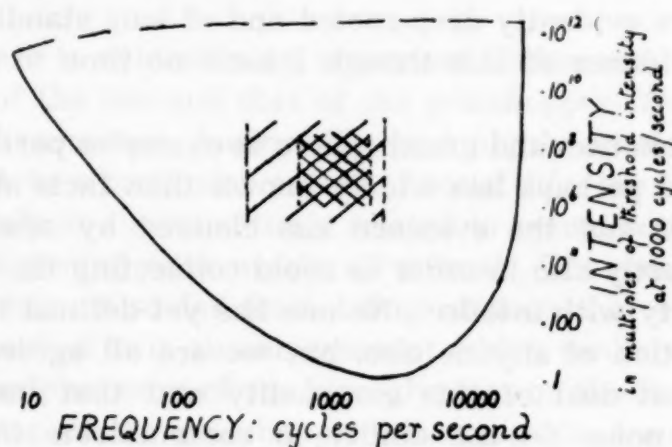


Fig. 2.

In fig. 2 the vertical and horizontal axes represent intensity and frequency of sound plotted on logarithmic scales. Each point on the diagram represents a tone of a particular intensity and frequency, and any complex sound can be represented by an appropriate constellation of points. Points lying within the circumscribed area correspond to audible tones; below it they are too faint to be heard and above it (here the limit is much less clearly defined) they are so loud that the ear is saturated. They may be painful, and accurate judgment of frequency, intensity or direction fails. It will be noted that the area is enormous, corresponding to about a thousandfold in frequency at its widest and a million millionfold in intensity at its deepest. Our

Post Office telephones pass a much more restricted band, roughly as shaded. And one type of phone passes sounds corresponding to the cross-hatched area and is still fairly intelligible. The Bell Telephone engineers set themselves to further economies. They found that all the intelligence could get through a system called a Vocoder which, instead of transmitting a continuous but limited spectrum, squeezed, as it were, all the sound energy of speech through ten narrow gates, thirty-two cycles wide as shown in fig. 3. We are not concerned with the economic consequence that,

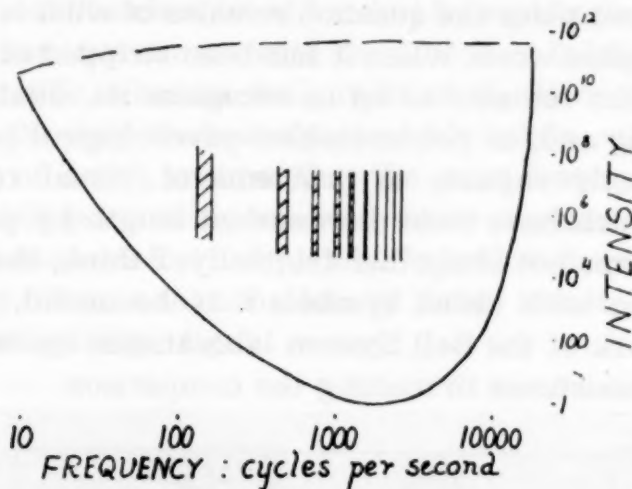


Fig. 3. Vocoder spectrum. On a linear scale the 'slits' would appear to be of equal width and spacing.

with sufficient paraphernalia at the sending and receiving ends, about ten intelligible messages can now be simultaneously transmitted over a channel where one would go before. The interesting feature for us is the effect of this process on the character of speech, for in discarding or blurring the detailed structure, the process has effected a complete mechanical separation of the emotive and informative functions of speech. The output of this rather infernal machine is perfectly intelligible and perfectly impersonal. No trace of anger or love, pity or terror, irony or sincerity can get through it. The age or sex of the speaker cannot be guessed. No dog would recognise his master's voice. In fact it does not sound as if a human agent was responsible for the message. But the intelligence is unimpaired. In fact its output conveys as much information as a typewritten transcript—and no more. Moreover it is not difficult to add a bogus personality and emotional content to the stripped message. It can be made

to sound like a soprano or a bass, flippant or impassioned, serious or silly, convincing or unconvincing.

Now, in the first place, this is a dramatic demonstration that our ears are not given to us simply to understand the intelligence of words. They are far too good for this job,—thousands of times too good. Any theory that human speech is denied to animals because they have neither ears to hear nor tongues to speak simply would not hold water. For if a far cruder ear would do, so would a far cruder mouth and larynx.

In the second place the question remains of what is the essence of an intelligible word. When it has been stripped of its emotive envelope, what remains to let us recognise its identity? This is an interesting and, as yet, untackled psychological problem, but it is manifestly cognate to problems of visual recognition—problems which have been discussed at length by psychologists though perhaps not altogether fruitfully. I think, therefore, that a comparison with visual symbols may be useful, particularly since the work of the Bell System laboratories on visible speech is of direct assistance in making the comparison.

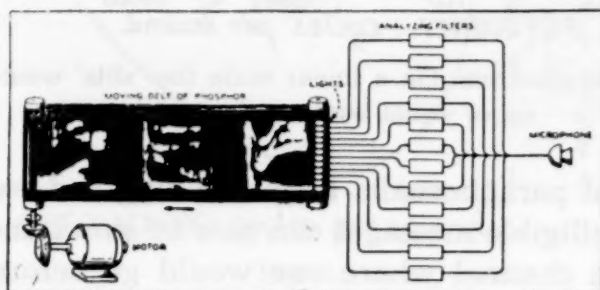


Fig. 4. Visible speech diagram. The words are "One, two, three". (Potter, Kopp, and Green, *Visible Speech*. D. Van Nostrand Co. Inc.)

The Bell Visible Speech system has features in common with the Vocoder but a totally different purpose. Its aim is to make speech intelligible to the totally deaf by transforming audible symbols into corresponding visible symbols. Fig. 4 makes the principle of operation clear. When words are spoken into the microphone on the right, the sound power is converted into electrical power with the same frequency distribution. Then the electrical power is sorted by filters into twelve separate channels. All frequencies in the upper half of the auditory range are discarded. The remainder are diverted into channels 300



cycles wide so that frequency components between 0 and 300 cycles/second enter the first channel counting from the bottom, between 300 and 600 the second and so on. The electrical power in each channel is then used to excite a filament lamp to a corresponding brightness and the lamp in turn excites a proportionate brightness in that part of a moving phosphorescent belt which is exposed to it. In this way spoken words are transformed into pictures which represent a continuous but very crude frequency analysis of the speech.

Now it should be clear what I meant by saying that this system has properties in common with the Vocoder. In this system the output is visible, whereas the output of the Vocoder is audible, but in both the very crude analysis which involves the discarding or blurring of detail, ensures that the gross features which contain the intelligence are preserved while the emotive envelope—the expression—is lost completely. Indeed in the case of Visible Speech it is obvious that the loss of detail results in a positive gain in intelligibility, fig. 5.

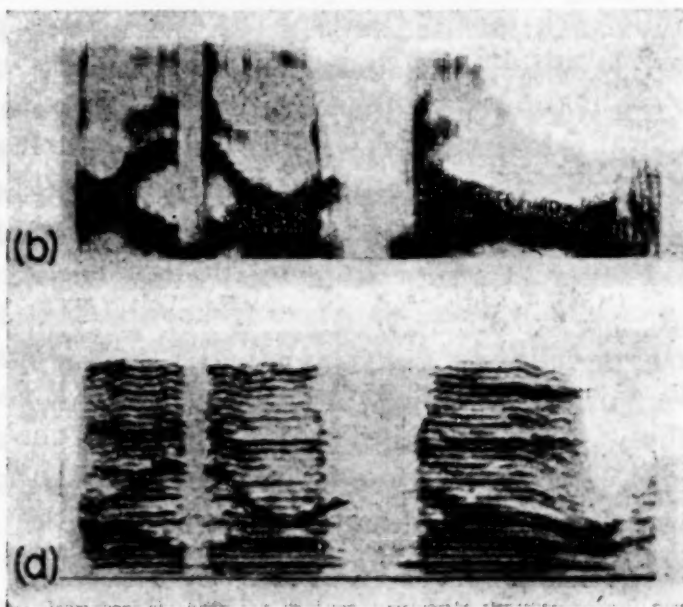


Fig. 5. The words are "I don't know".

(Above) ordinary Visible Speech system.

(Below) with an analyser of five times the resolving power and therefore five times as much information. The effect of the additional resolution is to conceal the "intelligence" while adding nothing to the information utilizable by the eye.

(Potter, Kopp, and Green, *Visible Speech*. D. Van Nostrand Co. Inc.)

These clear-cut results seem to me to flood the evolution of words with a new light. Let us briefly consider the probable evolution of the word signifying *lion*. Despite Max Müller's ironic strictures on what he called the "bow-wow" theory and the "yo-heave-ho" theory of the origin of words it seems very possible that the original sound symbol for lion was either imitative of a lion or expressive of the emotion appropriate to a lion's presence. In any case it would directly evoke in the people who heard it the emotions and behaviour suited to the actual presence of a lion. But as we have seen, those parts of the structure of a sound symbol which are emotive are not the parts on which ease of recognition and intelligibility depend. Provided that the brain had the capability of dealing separately with the intelligence and the rest of the information, then once the symbol had been invented it was certain rapidly to change in use in such a way that the intelligible structure was emphasized at the expense of the emotive. And this evolution must have had three consequences of profound importance for the further development of speech. Firstly, the detachment of the emotive envelope from the intelligible core meant that speech could be dispassionate. It was possible to discuss lions without looking over one's shoulder to see if they were there. Secondly, it meant an extension of temporal reference, for, whereas an emotive lion is necessarily in the present, an intelligible lion could be discussed in the future or in the past; and so tradition and forward planning about lions became possible. Thirdly, the dependence of intelligibility on a crude and easily appreciated sound structure meant that the symbol could be abbreviated in length. It takes much less time to say the word lion than to produce a recognisable imitation of the animal or to express adequately one's emotions on beholding it. So discussion could be enormously speeded up.

These three features, detachment, economy and an extension of future reference, seem to me to be the essentials by which human speech can be distinguished from that of animals. Of course the extension of temporal reference has been commented on before, but its dependence on the other two features has not been recognized and perhaps could not have been before the Bell Telephone work made a better understanding of the structure and evolution of words possible. These three clues now enable us, in my opinion, to put a finger accurately on the point

in history where human speech as we understand it began to be important.

It may well be at this point that some of you have been somewhat bewildered by technicalities. I should like, therefore, to try to illustrate the evolution of detachment and economy by means of another analogy, that of caricature. To avoid the possibility of offence to living celebrities let us consider the evolution of Colonel Blimp. (The work of the cartoonist, David Low may not be very familiar to Continental readers, but the argument applies, *mutatis mutandis*, to cartoonists generally, so I have not altered it.) Blimpishness has passed into the language. We are all familiar with the concept, and, though the original was probably composite, we may suppose him to have been an individual and we may further suppose that a good portrait of him in his Turkish Bath would illustrate not only the physical appearance we know so well but something of his character—his pomposity, his good-nature, his honesty, his obstinacy, his irascibility, his impenetrable stupidity. All these David Low represented with extraordinary skill in the greatly simplified outlines of the earlier Blimp sketches, outlines which gained in intelligibility because of the omission of detailed portraiture. But the process went much further until, for me at least, the whole concept of Blimpishness is symbolized by a pair of walrus moustaches, an abdomen with curious creases and the preposterous and inevitable bath-towel round the hips. The symbolic representation of Blimp has degenerated until there is no obvious correspondence between the ultimate symbol and the complex which it represents. It is, so to speak, no longer onomatopæic. It is very nearly an ideograph which is far more suited for purposes of communication than a portrait of the original would have been. Moreover, in addition to the economy of the degenerate final symbol, it has acquired detachment. It is impossible to feel for the symbol the dislike or affection which the original might have aroused. An emotional response can only be imported by retracing the evolution of the symbol, just as the word lion only induces an emotional response when one begins to think what it would be like to have a lion just behind one now (or when it is reclothed with emotion by being spoken in a terrifying way). The word lion and the bath-towel draped below a stomach protuberant and oddly creased are therefore degen-





starting from the time, roughly 600,000 years ago, when unquestioned implements of worked stone are found. In a recent, excellent summary of Stone Age cultures significantly, but misleadingly, entitled *Man the Tool Maker*, Oakley has chosen the same starting point; but there is little doubt that the beginning of fabrication of implements must be referred to an earlier date, perhaps more than a million years ago, and that the use of selected but unworked implements began much earlier still. On the grounds that the use of tools and the use of speech imply intellectual processes of the same order which are characteristically human, some authorities have inclined to refer the origin of speech to the lower Pleistocene or even to the Pliocene. There are, however, a number fallacies in this argument. Firstly, though anthropological evidence is extremely scanty, it is very doubtful if the Hominids of the Pliocene and lower Pleistocene are referable to the same genus as ourselves, certainly not to the same species. Secondly, there is no evidence that the earlier fabricated implements were *tools*, that is to say that they were made in order to *make* something else with them. And thirdly there seems to me no valid reason for assigning intellect to a maker of implements. I hesitate to import prejudice into the argument by using the contentious term *instinct*, so I will content myself with pointing out the obvious and incontrovertible fact that the web of a garden-spider and the nest of a chaffinch are highly fabricated implements quite as difficult to explain away as any product of lower Palaeolithic man. Certainly throughout the whole four or five hundred thousand years of the Lower Palaeolithic, the two main cultures—the hand-axe and flake cultures—show an extraordinary conservatism of type and an improvement in the technique of manufacture so gradual as to make the intervention of what we should call “reason” unlikely in the extreme. And so, *pace* Dr. Oakley, I feel that “Subman the Implement Maker” would have been a more accurate if less impressive title at least for the first half of his book.

And then in the last Ice Age the picture changes, at least in Europe, with dramatic suddenness. It used to be believed that at this point the autochthonous Neanderthal man was exterminated and replaced by an invader of much higher cultural achievement. But the facts seem to be much more consonant with a new stimulus and a sudden and rapid efflorescence than with the

total replament of an old and stable culture by a new and stable culture. This view is strengthened by anthropological evidence that the invading Cro-magnon type did not exterminate the Neanderthaler but interbred with him, although these races differ substantially in the form of the skull and were until recently assigned to different species, *sapiens* and *neanderthalensis* respectively. The direct evidence of interbreeding comes from the Near East, but, if it happened there, it could have happened wherever the invading Cro-magnon type, entering Europe from the East and from the South through North Africa and Spain, encountered the Neanderthalers held and pressed South by the advancing ice-sheets.

It is entirely in accordance with the principles of what, I suppose, we must now distinguish as Mendelist-Morganist genetics that such interbreeding between very distantly related stocks should induce very wide variability persisting through a large number of filial generations, and affecting both physical and mental traits. The stage is set for the production of both genius and imbecility with all their intergradations, with the probable consequences of specialization of occupation and the formation of an aristocracy of talent. If this background is speculative it seems at least plausible and consistent with efflorescence of manufacture and of art which did occur.

On the manufacturing side the beautifully finished flakes of the Neanderthalers are replaced by a wide range of sketchy but adequate flint tools—borers, scrapers, adzes, spokeshaves, shapers, polishers and graving tools of varying form and obviously of limited application. The conclusion seems inescapable, that in use they were applied successively in a sequence of operations, at first on wood but very soon on the more intractable bone and antler which has survived as evidence in the form of barbed harpoons, spear-throwers, and even fish-hooks and eyed needles. Here we have, for the first time, clear evidence of an objective reached through a planned and orderly succession of different operations.

On the artistic side there is the beginning of the very impressive series of cave paintings, but the technical difficulty of dating these or even of placing them in the right succession is very great I shall refer now only to a few objects, of which the dating is not in doubt and which are sufficient to illustrate my case.

Compare first the two engraved mammoths in fig. 7. In the first of Aurignacian age the outline is complete and the drawing is careful if rather clumsy. In the second of Magdalenian age there is a concentration on, and even an exaggeration of, the recognisable essentials, the outline of the head and back and the general impression of tusky woolliness. The eye is conventionalized and out of drawing. The legs and feet are not drawn at all but effectively suggested by incomplete ellipses.

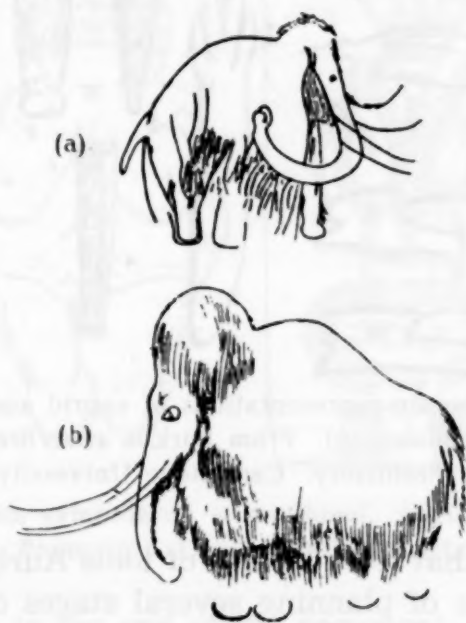


Fig. 7. Engraved mammoths.

(a) Aurignacian. (b) Magdalenian.

Burkitt, *Prehistory*. Cambridge University Press.

For reasons of which we are not yet sure, representations of animals are far more numerous in this age than representations of man, and realistic representation persisted far longer in this branch of art. But by the Magdalenian age degeneration even in the representation of animals was widespread. Figure 8 shows some examples of this age. The same trend is, however, discernible in representations of the human figure at a much earlier stage. In figs. 9 and 10 are examples, two of Aurignacian and one of Solutrean age. The first demonstrates that Aurignacian man could model a face if he wished and is included for comparison with the second, which represents not a woman but essential womanhood. The modelling is skilful enough but concentrates entirely on the recognizable female sexual charac-

teristics. The features which woman and man have in common, eyes, nose, mouth, ears and feet, are simply not there, or, like the arms and hands, barely suggested.

Finally in the Solutrean age the representation of womanhood has degenerated to a geometrical pattern.



Fig. 8. Degenerate representations of caprid and equine heads (Magdalenian). From Burkitt after Breuil.

Burkitt, *Prehistory*. Cambridge University Press.

We have seen that from his use of tools Aurignacian man was becoming capable of planning several stages of action into the

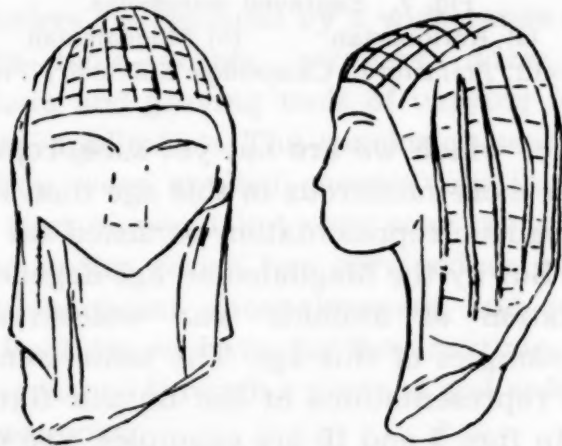


Fig. 9. Head of Woman (Aurignacian).

Burkitt, *Prehistory*. Cambridge University Press.

future. Here we seem to have equally unequivocal evidence that man had in this age reached a stage at which he could accept



degenerate symbols, not for what they were, but for what they had come to mean. In this era therefore, but not before, the three conditions, which I have previously stipulated, for flexible and characteristically human speech existed, and there is, so far as I can see, no evidence that they existed earlier.

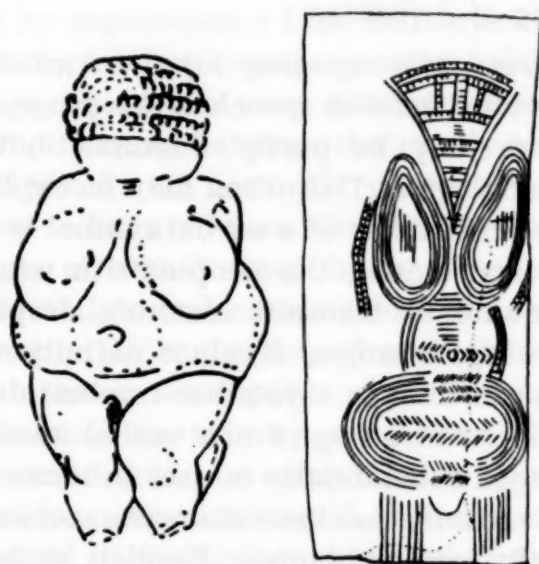


Fig. 10. Degenerate symbols of 'womanhood'. Aurignacian, Solutrean.  
Burkitt, *Prehistory*. Cambridge University Press.

This brings me to the end of the argument. I do not contend that my inferences are necessarily correct; and there may be errors of fact, though I have tried to avoid them. But I do assert that there is here a field which deserves further exploration by those who are competent to explore it.

#### SUPPLEMENT

I am deeply indebted to Professor Révész for offering to publish my lecture and for making a number of suggestions for improving the text, most of which I have adopted without hesitation. On two points, however, there remain outstanding differences between us which in different ways seem to be interesting and important and on which I do not feel disposed to yield. It seemed best to deal with these differences in a separate note. The text of the preceding paper therefore corresponds closely to my lecture save for the deletion of a considerable amount of matter which, however appropriate to the

original occasion, was only interesting and perhaps only intelligible to an Englishman and a member of my University. By refraining from patchwork I have, I hope, left it with as much unity and order as it ever had.

The points of difference are these:

### 1. *The meaning of definition*

Professor Révész, after quoting my own words (p. 1), states that I have nowhere defined *speech* or *language*. The difference between us here may be partly linguistic but I think it is primarily methodological. Definition may be explicit (formal) or implicit. Implicit definition of a verbal symbol is *its usage* in the field in which and among the people with whom that verbal symbol is adequate to transmit meaning. Implicit definitions change as the usage changes. Explicit definitions are attempts made from time to time to crystallise implicit definitions i.e. to express the current meaning of one verbal symbol (a meaning which has changed substantially in use) in terms of other verbal symbols whose meaning has been changing more slowly and which are consequently less ambiguous. Explicit definitions are like leaky lifebelts in the confused and angry sea of discourse; the rougher the sea the more irresistible the temptation to cling too long to them and the greater the ultimate probability of drowning in consequence.

Professor Révész is, I think, critical of the absence from my paper of *explicit* definition of speech and language. I maintain that the time is not ripe for explicit definition, because the usage, which is the implicit definition, is too protean. The whole paper is a contribution to the implicit definition of language; and the more the subject of language is discussed, the nearer the day when explicit definitions of at least temporary appropriateness can be made.

### 2. *The (explicit) definition of speech (language)*

Professor Révész contends that if I legitimately use *language* in a wide sense to include all the systems of communication among animals (including man), I must also admit "a special or strict meaning of *speech*, namely *human means of communication, (speech by sound, gesticulation and symbols (e.g. mathematical))*." This contention exemplifies a danger of premature

explicit definition in that it draws lines of distinction in the wrong places.

1. It fails to separate the two quite disparate components of human speech, the purely verbal component which is transcribable and the component of intonation and gesture which is not transcribable.

2. It makes by implication a false distinction between human and other speech, for *both* components are detectable in non-human languages though in very different proportions.

The essential difference between the purely verbal component and the intonational and gestural component is not that the former is more arbitrary than the latter. It is that words (and some signs or signals used by non-human animals) are unitary and particulate and each corresponds one-to-one with a field of meaning or with a group of such fields, whereas both gesture and intonation are capable of continuous variation and hence of expressing an infinity of shades of meaning. The ambiguity associated with one component is therefore of a quite different origin from that associated with the other.

It is by virtue of its particulate nature that the purely verbal component is transcribable; and the inventions of writing and printing have unquestionably helped to emphasize its importance in civilised communities to such an extent that the non-verbal component is largely discounted. The fond dog-owner who believes his pet to "understand every word" is obviously mistaken. But it is equally obvious that the dog extracts a great deal of information from human "speech" which passes unnoticed by a human auditor.

#### EPILOGUE FROM G. RÉVÉSZ

I do not wish to go into this question, phrased so clearly by Professor Pumphrey, as I have dealt with it quite elaborately in my book "*Origine et préhistoire du langage*", Payot, Paris 1950 (German by Francke, Bern 1946, English by Longmans Green, London 1954). Beside it will be treated again under the heading "Is there an animal language?", an article which will appear soon in Hibbert's Journal.

*Institute of Clinical Psychology University of Utrecht,  
The Netherlands*

## PROJECTION TESTS AND OVERT BEHAVIOR <sup>1)</sup>

BY

D. J. VAN LENNEP AND R. H. HOUWINK

### 1. THE PROBLEM

It is still a real problem if, and if so in what measure, projection tests give indications of the overt behavior of the subject. As Lindzey (9) has justly stated recently: "Available empirical evidence clearly indicates, that the assumed imperfect correlation between fantasied and overt behavior is warranted. However, at present we are far from an adequate formulation of the signs or cues that might permit specification from fantasy protocols alone of the behavioral tendencies that will secure overt expression as opposed to those that will not".

In practice a great number of possibilities present themselves: one time a projection test protocol gives fragments out of the earlier life's history which is no longer of actual interest, then again the content is a clear overcompensation of frustrated, unrealized, or in reality unrealizable wishes or ideals; sometimes the characters of friends or family are described, and then again the subject's own dynamic structure. Sometimes reflections of the actual situation or the day's activity appear, then again deeper lying structures, which the examiner thinks he can find in a projection protocol. This latter designation is, however, usually something that is only possible later when the protocol is compared with the detailed clinical and case-history data, and, however interesting this may be, we would rather have more certainty in the judgment of the test as such; in other words, a criterion with the help of which we could make out which of the different possibilities we have before us in a definite protocol we must judge.

The problem which we have outlined here is probably not only

---

<sup>1)</sup> We are much indebted to the Rockefeller Foundation, which has made this study possible by the sponsoring of research fellowships at the Institute of Clinical Psychology of Utrecht University.



a problem of interpretation, but also of test construction and test instructions. Perhaps up to now we have not been following the best method with our insistence upon entirely free-imagination projection tests. Symonds (14) comes to a very pertinent conclusion (p. 205): "If a person works out his problems in overt behavior, he does not find it necessary to work them out in fantasy—and if he works them out in fantasy, he is not bound to express them in reality. It is for this reason that seldom were the characters in a person's stories replicas of the person himself in real life". He then distinguishes different levels upon which fantasy and behavior function but he fails to tell us how we must decide on which level we should interpret the content of a single protocol.

In general projection tests up to now have been very disappointing where prediction over the subject's overt behavior is concerned, or rather we lack up to now fast criteria which can guide us in this respect since no one will deny, that *sometimes* overt behavior comes to expression in projection protocols, and *often* it doesn't.

Symonds does come to the conclusion that one must look for inner dynamics in the fantasy material from Thematic Apperception Tests, "that is, by assuming that they are projections of trends within the individual and then attempting to understand them in the light of all the possible transformations and disguises by which an individual protects himself from the anxiety which facing his unacceptable drives would arouse" (p. 209).

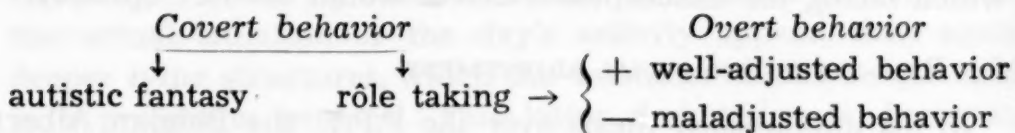
## 2. ROLE-TAKING AND SOCIAL ADJUSTMENT

In his unpublished thesis over the F.P.T. the Belgian Albert Collette makes the very just comment that the hero in the F.P.T.-protocols always behaves as an only child (1a). This is probably related to the fact that in our fantasy we are only concerned with ourselves; that the figures that live in our fantasy, such as those in our dreams, are also aspects of ourselves, or mostly nothing more than objects "to talk to", shadowy beings without the character of reality; and also that the assignment of rôles in our fantasy represents aspects of ourselves which furnish the dialectic necessary for all thinking, but which lack the degree of "adversité" that people from our own environment have, people from the hard reality, who help to determine our behavior

in the actual world. These figures out of our fantasy do not make "rôle-taking" necessary. They ask no empathy, no adjustment; without asking more, they satisfy our deepest tendencies and ideals if we wish; or they are the adequate correlates of our irrational anxieties and feelings of being menaced. However important this inner dialectic may be for the *interpretation* of overt behavior, an acquaintance with it does not always put us in a position to make *predictions* over overt behavior itself.

We can distinguish between two sorts of overt behavior: socially well-adjusted and socially maladjusted behavior. In so far as overt behavior is socially well-adjusted behavior, not individualistic behavior, but group-determined (or partner-determined) behavior is in the foreground. We can therefore say that overt behavior interests us, as clinical and industrial psychologists, in so far as it is not or not yet well-adjusted, i.e. in so far as it is still individualistic. Of course socially well-adjusted behavior remains "personal", but by individualistic behavior we mean here behavior in so far as it is under the influence of what Coutu (2) calls "the margin-of-error-in-rôle-taking" (p. 396). Besides, in the overt well-adjusted behavior the idiosyncrasies are, as it were, hidden behind, or controlled by the rôle-taking ability of the individual.

We can also say that the overt behavior interests us especially in so far as it is a function of the privative mode of rôle taking. We can make the following scheme:



Here we see that maladjusted overt behavior is not a direct function of autistic fantasy, but of a failure in rôle-taking, of the margin-of-error-in-rôle-taking.

### 3. ROLE TAKING IN THE F.P.T.

If there is any truth in this construction, we should be able to discern all sorts of things in relation to the maladjusted aspects of covert behavior by studying someone's *failure* in rôle-taking. We have tried to find out if this is possible via a projection test. In this case we must invite the subject, by special instructions, to assume a rôle-taking attitude instead of improvising freely over

a picture. For this purpose we used the Four Picture Test of Van Lennep (5, 6, 7). Instead of the usual instructions to make a story from the four given pictures, a number of arbitrarily chosen male subjects were asked to make a story with a female principal character <sup>2)</sup>. These instructions were chosen for several reasons.

From the statistical data available over the F.P.T. (Rümke, 12) we found that less than 1 % of our male subjects, when given the usual instructions, make a story spontaneously with a female principal figure. That this is only partly a function of the subject's sex appears from the fact that of the female subjects only 11 % made a story with a female principal figure. It seems therefore quite probable that the fact, that by far the most subjects make a story with a male principal figure, is also a function of the graphic stimulus of the F.P.T. In other words, the F.P.T., through nothing more than its graphic qualities, does not invite the spontaneous making of a story with a female principal figure. By giving the task of writing a story with a female principal figure, we force the subject to a psychic act which is relatively difficult and can only be solved by taking on a definite attitude, namely by transferring oneself to the rôle of a woman as hero of the story. It is perhaps possible to speak here of a true act of rôle-taking.

Actually what is asked of the subject is an act of "playing-at-a-rôle", in which implicit rôle-taking is implied. In his article: "Rôle-playing vs. rôle-taking" (3), Coutu says: "Playing-at involves an elementary form of rôle-taking, the verbalized fantasy by which the child learns how to take the rôle of another. Here the child both imaginatively and overtly pretends he is another person—not necessarily a particular other person, but often a stereotype of some functionary. Playing-at is "make believe" or "play acting". *Playing-at* a rôle is not necessarily *playing-at-being* another person; it may represent playing-at performing some socially prescribed function which he cannot actually perform since he cannot at his age occupy the appropriate social position; or it may even involve playing-at being a cow pony, machine-gun or airplane, with the appropriate vocalizations, sound or noises".

---

<sup>2)</sup> In this study, we shall refer to these protocols as "special" F. P. T.

Perhaps another explicite rôle would also have been possible as the task set in our test, but for male subjects a woman is the most extremely "other". By requiring of our male subjects a female hero, we required thus an act of playing-at-a-rôle, at least if they were willing to fulfil this task. Since the F.P.T. requires that four pictures must be combined into one story, it is necessary that the hero be given all sorts of emotional, volutive, and cognitive functions which determine the thread of the story. Therefore, pictures in which the rôle of the woman is more or less determined by the graphic meaning are less suitable for this purpose.

Our assumption was that subjects who were able to write a story with a woman as principal figure and in which feminine emotional relations etcetera really originated in this woman, would have an easier emotional contact with others than the subjects who were not able to do so.

We also investigated the possibility, that the F.P.T. protocols with these new instructions would contain more information over the (maladjusted) overt behavior of the subject. The aspect of overt behavior chosen to verify this assumption was dependence and submissiveness. It was assumed, that, if the woman in the "special" F.P.T. protocol was domineering, the writer, more often than not, would be dependent and submissive in his daily behavior.

#### 4. EXPERIMENTS

##### a. Method

The *material* for this study consisted of the written F.P.T. protocols of 311 male subjects who were examined by the

TABLE 1

Distribution of Population according to age	
Years of age	No. cases
15 — 20	86
21 — 25	71
26 — 30	74
31 — 35	45
36 — 40	15
41 — 45	10
46 & older	10
	<hr/> 311



"Netherlands Foundation for Industrial Psychology" (an institute for vocational guidance and personnel selection). By far the majority of these cases fell within the category of selection for special industrial functions. The educational level of the subjects varied from junior high school to university. The ages of the population were distributed as shown in the adjoining table.

The instructions given these cases were as follows. For those subjects who had already written the usual F.P.T. story earlier in the day (by far the majority in this case), the following was asked: "Do you still remember that you have had to write a story about four pictures?" This question was almost always answered in the affirmative. The instructions continued then: "Make a story now about the same four pictures, in which one of the ladies at the tennis court plays the main rôle" <sup>3</sup>).

If the subject could not remember the four pictures, he was permitted to see them again. In those cases in which no ordinary F.P.T. had been done, the above instructions were added to the usual instruction. Whether or not the "special" F.P.T. was preceded by the usual F.P.T. proved to have no influence on the results.

#### b. *Analysis of the Results*

From further study of the protocols collected in the manner described above it appeared that three groups of stories could be clearly differentiated, namely:

1. Stories in which a woman was consistently the principal figure, where she determined the course of action and gave the story a specifically feminine content (henceforth called "feminine stories" for short).

2. Stories in which a man was really consistently the principal figure while a woman appeared only as an extremely secondary figure or not at all (so-called "masculine stories").

3. Mixed stories in which it was not certain whether a man or woman was the principal figure, either because there was no really principal rôle, or because a man and a woman appeared as perfectly equal figures. For criteria see below.

These story-types were then compared with the subject's

---

<sup>3</sup>) The fourth picture of the F. P. T. depicts a tennis court with players and spectators, two of whom are girls.

"affective contact" by which was understood his overt behavior during the examination in respect to the examiner, as the latter described this in his report over the examination. Thus the term "affective contact" as used here does not imply prediction, but is based purely upon observed behavior in respect to the examiners. Although "easy affective contact" is not exactly the same as "good social adjustment", we prefer to use the former term, since it points to behavior observable in the examination, whereas the latter is more of an inference.

It proved possible to establish fairly objective, usable directions for the rating of the stories. In general, the following standards were used as the basis of differentiation:

*Feminine stories* were characteristically those in which a female principal figure appeared consistently through the whole story, where the thoughts, adventures, feelings and experiences of the woman were more or less explicitly described, unless these experiences were named exclusively as being negative in respect to an actual male figure in the principal rôle and were responsible for his behavior. At the same time, the woman was mostly the center in the story, played a specifically feminine rôle within its framework, and was directly, actively involved in the development of events. It is striking, that Happy-End stories were mostly feminine stories.

In *masculine stories*, on the other hand, the female figure either did not appear at all, was present only in the beginning, or remained entirely in the background. Often only events were described which followed each other in a simple action plot, even though a woman played a secondary rôle in this plot. Often the female figure appeared only as the center of the thoughts of one or more men (for example: a boy dreams about the girl next door; or: two boys in love with the same girl quarrel over her at the tennis court). Also judged as masculine stories are those in which a woman is only causally or indirectly concerned in the plot, and finally also those stories which actually have a woman as principal character, but in which her rôle is explicitly of the sort that it could be played just as well by a male figure and would be even more acceptable in this form. It is noteworthy that Detective stories were almost always masculine stories.

A few examples may clarify these standards of judgment for feminine and masculine stories.

## EXAMPLES

A. *Feminine Story*

(Subj. 75856, male, 29yrs)

Marian had put the children to bed and wanted to tidy up a little in the bedroom where she and her husband slept. On entering she paused: the moonlight lent the spacious room something mysterious and also something hidden. She had a warm feeling of "my home is my castle" at which she yet smiled: oh well, let's be deliciously sentimental for once! Well, yes, being sentimental is of course nonsense, but she was a lucky person and she had a cosy evening in front of her too, for downstairs at the moment sat Charles with Tom his friend deep in talk, and in a moment she would go and sit with them, making a remark now and then, but chiefly listening and.... enjoying herself. Her friends did not envy her for it, but she herself knew, that she was fond of evenings like this. They had got to know Tom last summer at the tennisclub, and that was how the friendship had begun. Just as she was drawing the curtains in the bedroom her eye fell on the opposite side of the street: it had begun to snow softly and a man stood with his collar turned up: old, lonely, poor. She was painfully touched, drew the curtains carefully and went pensively downstairs.

B. *Masculine story*

(Subj. 75924, male, 20yrs)

*An Accident*

At a tennis club consisting of three ladies and three gentlemen there was an accident one day while a game was being played. To wit: one of the men who was playing with a lady against another lady and gentleman at a certain moment hit — by accident — the lady on the head with his tennis racket. The blow came down hard, and the lady was taken to the hospital. The young man to whom this happened is beside himself with anxiety. At night he cannot sleep, lies for hours staring at the ceiling and the strangest thoughts go through his head. That she will die, that she will be crippled for life, etc. Because of this he is unbearable at home. At the least thing that is said to him he flies up and snarls at everyone. At last he gets news — a week later — that the girl is recovering, and that he may go to visit her. Long before it is time for the visit he is waiting in front of the hospital. To make things worse it begins to rain. But he doesn't even feel it. Without a coat he stands there in the rain till the right time has come. He is madly happy!

In order to see wheter the judgment of stories according to the above mentioned standards would remain at all constant, a small sample of the total material was given to someone who had no further connection at all with the study, with the request to judge those stories acording to the given criteria. Indeed, by this method, 82 % of the stories were judged in entirely the same

manner, while the remaining 18 % consisted only of those stories which by one judge were placed in the masculine or feminine category, and by the other in the "mixed" group. Judgements of the same stories by the same judge after a period of six months rendered almost complete agreement.

### c. Results

In the 311 cases studied here it was indeed found that the "feminine stories" came preponderantly from subjects who had a good and easy affective contact, while the opposite proved to be true of the writers of "masculine stories".

After applying the *Yates* correction formula,  $X^2 = 34.6$ . The Null Hypothesis can thus be discarded at the 0.1 % level of confidence. Our first hypothesis is therefore confirmed by the experimental results.

We also investigated the possibility that in respect to age and educational level of the subjects differences might be found between the groups which would be related in any way to the criterion. This, however, proved not to be the case.

Our second assumption was that the protocols of the "special" F.P.T., especially of the so-called "masculine stories", should contain more data over overt behavior than those protocols written with the usual task. There are at the present time a great many indications that this is really the case, but research over this problem is still in progress, and we hope to be able to publish more detailed information on this subject later. Up to the present time we have studied in detail one of the aspects of overt behavior which comes to expression in the "special" F.P.T. This aspect is submissive, dependent behavior (lack of independence) in subjects which we have brought into relation with the degree of dominance of the woman appearing in the "special" F.P.T. The independence of the subjects was determined by us upon the basis of the reports of the examiners. In most of our cases it was explicitly stated by the examiner whether the subject made a dependent or independent impression, since this was considered an important factor in selection for industrial functions. These judgments were all clinical observations since there are no objective tests for this variable available here.

The supposition that dependent subjects write a "special"



F.P.T. in which the female figure plays a strongly dominant rôle, proved indeed to be true in our material, especially in the group of so-called masculine stories.

This fact is apparent in Table 2, which gives  $X^2$  and  $P$  for the relationship between these factors. When we consider the data over masculine and feminine stories apart (the groups "dominant  $\pm$ " and "dominant  $-$ " were considered together here, since otherwise the numbers would have been too small for statistical computations), it appears that, as we expected upon the basis of our hypothesis, in the masculine stories the degree of dependence in overt behavior was represented by the appearance of a more or less dominant female figure in the "special" F.P.T. story, while this proved not to be the case for the feminine stories.

TABLE 2

Independence of  $S$ , and dominance of woman in "special" F. P. T.-protocol  
( $N = 259$ )

	$X^2$	$P$
Whole Material:	31.45	$< 0.1 \%$
"Masculine Stories" only:	16.51	$< 0.1 \%$
"Feminine Stories" only:	1.29	$\pm 65 \%$

## 5. DISCUSSION

From the fact that our hypotheses have been proven true by this study we shall be able to draw a few general conclusions in respect to the application of projection tests. By permitting a subject to be entirely free in exercising his imagination over a projection picture, we have not made it necessary for him to assume a rôle-taking attitude, but have allowed him to express himself within the dialectic of his more autistic train of thought. Up to now we lack criteria for interpreting these autistic trains of thought into terms of overt behavior. However, it is possible to present tasks in which the subject is forced to place himself in the rôle-taking attitude. For male subjects it is a question of playing-at-a-rôle, to think and to act from the standpoint of a female principal figure. The F.P.T. is a good stimulus for this, since less than 1 % of the male subjects spontaneously make stories with a female principal figure. For female subjects it is naturally another question. In order to study the same phenomena in female subjects, a feminine F.P.T. has been con-

structed<sup>4)</sup>, in which most female subjects make stories with a female principal figure, when given the usual, ordinary instructions. Female subjects have been studied upon the basis of this test with the instructions to make a story with a male principal figure.

These investigations have not been finished yet, but on the basis of a small material we found, that here too aspects of the overt behavior were significantly related to variables in the protocols. These findings will be reported more in detail in a later publication. We are of the opinion, that large-scale research should be done concerning the question whether projection tests with instructions which imply rôle-taking, procure more information over overt behavior than those given with the free instructions used up to the present time.

Meanwhile, it is interesting, that Sarbin and Farberow (13) have found in their experiments on the subject of age-regression in relation to self and rôle, that there is a highly probable relation between social adjustment and rôle-taking abilities.

Gough (4) has very justly remarked, that "the basis for individual sociality is social interaction, and this interaction is effective in so far as the individual can look upon himself as an object or can assume various rôles", while Cameron (1, p. 93) states: "the more effectively (an individual) is able to allow the attitudes and responses of others, which he predicts in symbolic rôle-taking, to influence his own reactions, the more competent he ought to be in social situation". These various statements lend considerable support to our point of view.

We admit that our criterion: the subjective judgments of the examiners during a psychological examination, is not very satisfactory one from an objective point of view, but in Holland we have no "objective" instruments for the measurement of

---

<sup>4)</sup> This F.F.P.T. (Female Four Picture Test) consists of the following four drawings:

- I. Sitting room. A woman is bending over a cradle in which a baby can be seen. A small boy is sitting on the floor, playing with toys.
- II. Bedroom with old-fashioned furniture and a mirror.
- III. Lake in forest scenery where a female figure stand pondering.
- IV. Office. A male and a female figure are seated at desks talking with each other, while they are being watched by a second male figure from the background.

emotional adjustment. We suggest that this study might be duplicated in America where such instruments are available.

Also it remains to be seen, whether our findings are culturally determined and whether, if duplicated in America, the results would be similar. This is of special importance in so far as the results with the F.F.P.T. are concerned, because the culturally determined position of woman is so different in America and in Europe.

And it might be of interest to investigate how the different culturally determined characters, as Riesman worked them out in "The Lonely Crowd" (11) behave in regard to rôle-taking in the F.P.T. protocols with special instructions.

## 6. SUMMARY

In this article we have discussed the problem of the lack of sufficient criteria to permit a judgement on the basis of projection test protocols concerning whether a definite content is related to the subject's overt behavior.

Two sorts of overt behavior have been distinguished—socially well-adjusted behavior and socially maladjusted behavior. The former was brought into relation with general ability at rôle-taking, the latter with margin-of-error-in-rôle-taking (Coutu).

By using special instructions for the Four Picture Test (namely, making a story with a female principal figure), male subjects were forced to use their rôle-taking ability.

For our population there is evidence that:

1. Males with an easy affective contact, in so far as observed by the examiners of the "Netherlands Foundation for Industrial Psychology", can complete this task significantly better than males with a poor affective contact.

2. For the latter, the appearance of a dominant woman in the special F.P.T. protocol is significantly related to a submissive, dependent over behavior, as far as observed during the psychological examination.

There are indications, that still many other features of overt behavior appear in the stories of the unsuccessful rôle-takers. We hope to publish more material concerning this question later.

## RÉSUMÉ

Le présent article traite l'absence des critères qui permettraient de juger, en partant du protocole d'un test de projection, s'il existe un rapport entre le contenu déterminé d'un récit et la conduite manifeste du sujet.

On distingue deux formes de comportement manifeste: le comportement social adapté et le comportement social mal adapté. On rattache le premier à une aptitude générale pour „rôle-taking” (la faculté de se mettre à la place d'un autre); le second à la „margin-of-error-in-rôle-taking” (Coutu) (le degré d'insuccès en se mettant à la place d'un autre).

En exigeant des sujets d'écrire une histoire dont le personnage principal sera une femme ou une jeune fille, on a obligé des sujets du sexe masculin à exercer leur faculté de „rôle-taking”. L'examen du matériel obtenu a démontré:

1. que les hommes dont le contact affectif est aisé s'acquittent remarquablement mieux de cette tâche que les hommes avec qui un contact affectif s'établit difficilement, pour autant que le fait a pu être constaté par les examinateurs de la „Fondation Hollandaise pour la Psychotechnique”;

2. que chez les sujets appartenant à la seconde catégorie, l'apparition d'une femme dominatrice dans le protocole à l'instruction précité, correspond significativement à un comportement manifeste influençable et dépendant, pour autant qu'il a été constaté au cours de l'examen psychologique.

Il est dénoté que nombre d'autres aspects du comportement manifeste se révèlent dans les histoires écrites par les sujets qui ne parviennent pas à exécuter cette tâche de „rôle-taking”. Les auteurs espèrent publier ultérieurement des données plus élaborées touchant cette question.



## REFERENCES

1. Cameron, Norman: *The Psychology of Behavior Disorders*. Boston: Houghton Mifflin Co., 1947.
- 1a. Collette, A.: *Etude de la fiction chez l'adolescent par le Four Picture Test de D. J. van Lennep*. 2e Licence en sciences pédagogiques, Université de Liège, 1951.
2. Coutu, W.: *Emergent Human Nature*. New York: Alfred A. Knopf, 1949.
3. ———: *Rôle-playing vs. rôle-taking: An appeal for clarification*. *Amer. Sociol. Rev.* 1951, 16: 180—187.
4. Gough, H. G.: *A sociological theory of Psychopathy*. *Amer. Journ. Sociol.* 53, 359—366.
5. Lennep, D. J. van: *Manual Four Picture Test*. Martinus Nijhoff, The Hague, Holland, 1948.
6. ———: *The Four Picture Test*; in: Anderson & Anderson: *An Introduction to Projective Techniques*. New York: Prentice-Hall, 1951.
7. ———: *Psychologie van Projectieverschijnselen*. Utrecht: N. S. v. P., 1948.
8. ———: *De l'importance de la structure de la personnalité dans le contact affectif*. In: *The Affective Contact, Proceedings of the International Congress for Psychotherapeutics at Leyden 1951*. Amsterdam: Strengholt, 1952.
9. Lindzey, Gardner: *Thematic Apperception Test: Interpretative Assumptions and Related Empirical Evidence*. *Psychol. Bull.* 1952, 49: 1—25.
10. Mead, G. H.: *Mind, Self and Society*. E. W. Morris (Ed.). Chicago: The University of Chicago Press. 1939.
11. Riesmann, D.: *The Lonely Crowd*. New Haven; Yale University Press, 1949.
12. Rümke, A. C.: *Resultaten van een statistisch onderzoek van een aantal variabelen in 700 schriftelijke vierplaten-test protocollen* (Results of a statistical analysis of some variables in 700 written four-picture-test protocols). *Nederl. Tijdschr. Psychol.* 1952, 7: 337—363.
13. Sarbin, Th. R. and Farberow, N. L.: *Contributions to rôle-taking theory: A clinical study of self and role*. *Journ. Abnorm. Soc. Psychol.* 1952, 47, 117—125.
14. Symonds, P. M.: *Adolescent Fantasy*. New York: Columbia University Press, 1949.

## EDGAR RUBIN

Etwa vor zwei Jahren ist mein alter Freund Edgar Rubin, Professor der Psychologie an der Universität zu Kopenhagen, gestorben. Meine Absicht war über ihn in der *Acta Psychologica* ein Nekrolog zu schreiben, seine wissenschaftliche Persönlichkeit und seine psychologische Forschungsarbeit darzustellen. Ausserdem wollte ich den Einfluss dieses originellen und selbständigen Psychologen auf unsere Wissenschaft schildern. Zu diesem Zwecke habe ich mich bemüht, biographische Daten und Einzelheiten über die Tätigkeit Rubins als Lehrer und Organisator und die Liste seiner Publikationen aus seiner Heimat zu verschaffen. Das gelang mir leider bisher nicht. Daher muss ich mich damit begnügen, den Tod dieses hervorragenden Menschen und Gelehrten unseren Lesern mitzuteilen. Ich hoffe, dass unter meinen Kollegen in Skandinavien bald sich jemand finden wird, der für unsere Zeitschrift, deren ständiger Mitarbeiter er war, und in der seine letzte ausserordentlich bemerkenswerte Arbeit veröffentlicht wurde (*Visual figures apparently incompatible with geometry*, Vol. VII, p. 365), eine ausführliche Darstellung seiner wissenschaftlichen Tätigkeit geben wird.

G. RÉVÉSZ

## PSYCHOLOGIE DES SICHERHEITSMARGINALS

VON

DAVID KATZ

Wer einen internationalen Kongress zu organisieren hat, dem bieten sich mannigfache Gelegenheiten zum Studium seiner Kollegen. So ist es z.B. nicht ohne Interesse, die Reaktionszeiten der Kongressteilnehmer auf die zur Versendung kommenden Rundschreiben zu studieren. Manche zeigen ein Verhalten, das man als „vorzeitige Reaktion“ bezeichnen könnte. Sie melden nämlich Vorträge an auf das blossе Gerücht hin, dass ein Kongress stattfinden soll. Andere Kollegen haben eine positive, aber kurze Reaktionszeit. Sie antworten postwendend auf jedes an sie gerichtete Schreiben. Es ist unnötig zu sagen, dass diese Kollegen sich die besondere Wertschätzung des Organisationskomitees erwerben. Aber ebenso hat jeder Kongress auch seine Problemkinder. Das sind diejenigen Kollegen, die das Komitee am längsten mit Antworten warten lassen oder überhaupt nicht antworten, sodass dieses nicht nur in Schwierigkeiten bei der Aufstellung des Arbeitsprogrammes gerät, sondern auch wenn es weniger wichtige, wenn auch nicht ganz unwichtige Dinge gilt. Es ergibt sich hiernach differentialpsychologisch eine Dreigliederung der Kongressteilnehmer, eine Kategorie mit gut angepasstem und zwei mit weniger zweckmässigem Verhalten. Die Späten setzen sich der Gefahr aus zu spät zu kommen, die Vorzeitigen machen sich und dem Organisationskomitee unnötige Mühe. Wenn wir uns des Begriffes des Sicherheitsmarginals bedienen, so können wir sagen, dass die Frühen mit einem zu grossem, die Späten mit einem zu kleinen und die übrigen mit einem angemessenen Sicherheitsmarginal arbeiten.

Der Begriff Sicherheitsmarginal oder Sicherheitsfaktor ist der Technik entnommen. Seine Uebertragung auf die Psychologie erweist sich nicht nur als möglich, sondern auch als sehr förderlich für neue Fragestellungen. Dieses an einer Reihe von Tatsachen nachzuweisen, habe ich mir als Aufgabe meines

Vortrages gesetzt. Mein Thema lautet also: Psychologie des Sicherheitsmarginals. Das Sicherheitsmarginal soll nicht nur als ein technisches Detail unseres Sicherungsverhaltens behandelt, sondern soll in den grösseren Zusammenhang der Bedürfnispsychologie eingeordnet werden, in den es gehört, nämlich in den Zusammenhang mit der Frage des menschlichen Verhaltens, das der Erhaltung und Sicherung unserer Existenz dient. Dieses Verhalten ist im Gegensatz zu der nach rückwärts gerichteten Gedächtnisarbeit der Zukunft zugewandt.

Als unverbesserlicher Experimentalpsychologe kann ich es mir nicht versagen, durch ein Demonstrationsexperiment zu verdeutlichen, was unter „psychologischem Sicherheitsmarginal“ zu verstehen ist. Es ist sehr einfach und besteht darin, dass ich ein mit Wasser gefülltes Glas hebe. Dies gelingt nur, wenn ich mit den Fingern einen so grossen Druck auf das Glas ausübe, dass die Haftreibung ein Gleiten des Glases aus der Hand verhindert. Wie gross ist dieser Seitendruck? Er variiert beträchtlich individuell. Aber er ist immer wesentlich grösser als der minimale Druck, bei dem das Glas gehoben werden kann, ohne zu rutschen. Man kann den Seitendruck etwa unter Ausnutzung der Piezoelektrizität bestimmen, wobei Druck in elektrischen Strom umgesetzt wird <sup>1)</sup>. Und so kann man auch die Differenz ermitteln, die zwischen dem minimalen Druck und dem wirklich angewandten beim Heben des Glases besteht. Es ist diese Differenz, die wir für den Fall der Gewichtshebung als das individuelle Sicherheitsmarginal bezeichnen wollen <sup>2)</sup>. Den Erfahrungen bei der Gewichtshebung — eine seit G. Th. Fechner schon nahezu sakramental gewordene Angelegenheit der experimentellen Psychologie — gewinnen wir auf diese Weise eine neue Seite ab. Sie verdient unser Interesse um so mehr, als der dabei einsichtig gewordene Sachverhalt zu einer generalisierbaren Definition führt. „Das individuelle Sicherheitsmarginal bei jeder beliebigen Leistung ist der Aufwand an Arbeit, den jemand über das Minimum hinaus leistet, um die auszuführende Operation sicherzustellen.“ Das Problem des Sicherheitsmargi-

<sup>1)</sup> F. I. I. Buytendijk, Die Abstufung der willkürlichen Muskelspannung. Commentations, Vol. IV Nr 13. Pontifica academia scientiarum 1940.

<sup>2)</sup> Vgl. hierzu D. Katz und G. Korjus, Muskeltonus der Hand und Sicherheitsmarginal. Acta Paediatrica, Vol, 31, 1944.



nals gehört in erster Linie in die allgemeine Psychologie, doch bietet es darüber hinaus mancherlei spezielle Aspekte, z.B. einen differentialpsychologischen. Es ist kein Zufall, wenn manche Individuen zu einem grossen Marginal neigen, andere zu einem kleinen, wenn manche eine grosse, andere eine kleine intra-individuelle Variation zeigen. Für Typenjäger eröffnet sich hier ein neues Feld der Betätigung. Der differentialpsychologische Gesichtspunkt führt leicht zu einem mentalhygienischen und psychotherapeutischen. Es ist unnötige Vergeudung von psychophysischer Energie, wenn man das Marginal zu gross nimmt, es ist riskant, wenn man es zu klein nimmt. Auch auf die Psychologie des Unfällers fällt von hier aus neues Licht.

Der Begriff des Sicherheitsmarginals lädt zu neuartigen sozialpsychologischen und völkerpsychologischen Untersuchungen ein. Wie sich hinsichtlich der Neigung zu sparen die verschiedenen sozialen Schichten voneinander unterscheiden, so gilt dies auch völkerpsychologisch. Unsere englischen Freunde wissen ein Lied von der grauen „austerity“ zu singen, die gegenwärtig ihr nationales Leben beherrscht. Andere Völker in ähnlicher ökonomischer Lage haben nicht die gleichen Konsequenzen gezogen für die Ordnung der Volkswirtschaft. Der internationale Kongress in Edinburg schloss mit einem gewaltigen finanziellen Ueberschuss, ich kann dem Präsidenten des nächsten Kongresses in dieser Hinsicht keine grossen Aussichten machen. Schweden ist nicht Schottland.

Bevor wir uns eingehender den Fragen des Sicherheitsmarginals innerhalb der Psychologie zuwenden, empfiehlt es sich, die Frage der Sicherung von Leistungen auf zwei anderen Gebieten zu berühren, und zwar einerseits innerhalb der Technik, andererseits innerhalb der Physiologie, wo unter anderen Cannon im Zusammenhang mit seinen Untersuchungen zur Homöostasis Entscheidendes zur Lösung des Problem der Sicherung des Lebens und damit zur Physiologie des Sicherheitsmarginals beigetragen hat <sup>3)</sup>.

In jedem schwedischen Lift findet man einen Anschlag darüber, wieviel Personen dieser zu befördern vermag. Es heisst dort etwa, dass höchstens eine Person zulässig ist. Aber was

---

<sup>3)</sup> Walter B. Cannon, *The Wisdom of the body*. Second Edition. London 1948.

heisst hier eine Person? Zwei in Hollywood zugelassene Filmsterne wiegen wahrscheinlich nicht so viel wie der Mann, der die Weltmeisterschaft im Gewichtsstemmen innehat. Aber der Lift würde vermutlich nicht einmal seinen Dienst versagen, wenn man ihm zwei dieser Muskelkolosse anvertrauen würde. Mit noch weit grösserer Sicherheit als beim Lift wird bei anderen Gelegenheiten gerechnet. Was die Konstruktion von Brücken, Dammen, Wasserreservoirs und Gebäuden angeht, so folgt man hier häufig der sogenannten Bach'schen Regel, die vierfache Sicherheit gegen Bruch verlangt, in der Bedeutung, dass das Verhältnis zwischen den Werten, die der Bruchgrenze des Materials und denjenigen der maximal zugelassenen Belastung entsprechen, wenigstens vier sein sollte<sup>4)</sup>.

Seltsamerweise gibt es keine allgemeine Definition des Begriffes des Sicherheitsfaktors in der Technik. Die Vorschriften, die man von Fall zu Fall aufstellt, gehen nicht von einem gemeinsamen Prinzip aus und stützen sich nicht auf eine klare Auffassung vom Inhalt des Sicherheitsbegriffes oder seiner Funktion. Erst in der modernen Flugtechnik, wo die Forderung der Gewichtsökonomie eine Revision des Sicherheitsbegriffs nötig machte, ist man zu einer mehr wissenschaftlichen Behandlung gekommen. Man kann wohl sagen, dass in der Technik alle möglichen Sicherungen innerhalb der Grenzen 400 und 1500 % vorkommen. Ich habe absichtlich bezüglich technischer Einrichtungen von Sicherheitsfaktoren, nicht aber von Sicherheitsmarginal gesprochen. Die technische Sicherheit kommt zum Ausdruck durch den Abstand zwischen der aktuellen Inanspruchnahme einer Einrichtung und der Bruchgrenze oder — wenn es sich um Maschinen handelt, — der oberen Leistungsgrenze. Jede technische Einrichtung reagiert aber, wie gross auch ihre Reserven sind, stets mit dem Wert, der von ihr „hic et nunc“ gefordert wird, niemals mit der kleinsten Zulage darüber. Brücken, Dämme, Häuser sind technische Gebilde *statischer* Natur, die Erwägungen über die Sicherheit bedürfen einer Ergänzung, wenn es sich um Maschinen handelt, die der jeweiligen Inanspruchnahme durch menschlichen Eingriff oder nach kybernetischen Prinzipien automatisch angepasst werden. Aber auch

---

<sup>4)</sup> Vgl. hierzu Folke K. G. Odquist, Festigkeitslehre. (Schwedisch.) Stockholm 1948.

bei ihnen gibt es keine Analogie zum Sicherheitsmarginal im eingangs definierten Sinne, das eine spezifisch bio-psychologische Einrichtung ist. Hebe ich Gewichte, so nimmt das Sicherheitsmarginal mit wachsenden Gewichten dauernd ab, bis es beim grössten Gewicht, das ich gerade noch bewältigen kann, den Nullwert tangiert. Je grösser die Steigung, die ein Automobil zu überwinden hat, um so kleiner wird die Kraftreserve, um bei der grössten noch überwindbaren gleich Null zu werden, aber von Anpassung des Sicherheitsmarginals kann man hier sowie bei allen anderen technischen Vorgängen nicht sprechen. Das hängt unter anderem auch damit zusammen, dass nur das menschliche Handeln, aber nicht das maschinelle Geschehen affektiven Einflüssen unterworfen ist. Eine Lokomotive benimmt sich nicht anders, ob sie mit Kohle oder mit Dynamit geladene Wagen zu ziehen hat, wohl aber legt sich bei der gefährlicheren Last die Hand des Lokomotivführers fester um den Griff des Regulators der Lokomotive. Wenn man jemanden zwei Pakete von demselben Gewicht und derselben äusseren Form tragen lässt, von denen das eine Seife enthält, das andere Eier, so schliessen sich die Finger um das letztere kräftiger. Eine Hausfrau, die auf der Mottenjagd ist, tötet das verhasste Insekt mit einem Schlag, der für hundert Motten ausreichend gewesen wäre.

Wie steht es mit dem Sicherheitsproblem in der Physiologie? Der menschliche Organismus ist so konstruiert, dass er wie eine Maschine über nicht unbeträchtliche Reserven verfügt. Von paarig vertretenen Sinnesorganen — Auge und Ohr — kann das eine ausfallen, ohne dass der Mensch dadurch ernsthaft in seiner Existenzfähigkeit bedroht wird. Der Mensch ist auch mit einer Niere lebensfähig; er kann mit einem Lungenflügel leben, wenn der andere durch Pneumothorax ausser Funktion gesetzt ist. Gibt es auch physiologische Einrichtungen und Funktionen, die im Sinne des früher definierten Sicherheitsmarginals arbeiten? Für einige kann man das hypothetisch annehmen, für andere kann man es als erwiesen ansehen. Es ist zum mindesten wahrscheinlich, dass solche Vorgänge wie die Beschleunigung der Herztätigkeit bei stärkerer Inanspruchnahme durch Arbeit, dass die Adrenalinabsonderung bei emotionaler Umstimmung und die Mobilisierung der Leukozyten gegen bakterielle Eindringlinge über das Ziel und individuell verschieden über das Ziel hinausschiessen. Was die beiden natürlichen Zangen angeht,



mit denen wir von Natur ausgerüstet sind, die Handzange und die Kieferzange, so arbeiten diese von Geburt an im „clinging reflex“, resp. bei der Umklammerung der Brustwarze mit einem sehr beträchtlichen Sicherheitsmarginal.

So viel über das Sicherheitsproblem in Technik und Physiologie. Lassen Sie uns zu den spezifisch psychologischen Problemen des Sicherheitsmarginals zurückkehren und versuchen diese zu klassifizieren. Es sind Handlungen im gegenwärtigen Raum und in der gegenwärtigen Zeit, auf die unser Begriff in erster Linie in Anwendung kommen kann. Diese Handlungen werden meist routinemässig vollzogen und haben als solche nichts Aufregendes. Daneben gibt es aber dramatische Augenblicke im Leben, bei einer Explosion oder Feuerbrunst, wo blitzschnelle Entschlüsse gefasst und ausgeführt werden müssen und wo das Sicherheitsmarginal wohl immer die Nullgrenze berührt. Alle ausserhalb des aktuellen Raum-Zeit-Rahmens liegenden, auf die Sicherung ausgehenden Leistungen stützen sich mehr oder weniger auf bewusst verlaufende Reflexionen über die Zukunft, über das Morgen, das nächste Jahr, das ganze Leben, ja sie langen mit dem Testament über unsere Lebenszeit hinaus.

Das Heben eines Gewichtes an dem ich einleitend das Wesen des Sicherheitsmarginals illustriert habe, ist ja ein trivialer Arbeitsvorgang. Etwas interessanteres ist schon der Schreibakt, bei dem das Schreibinstrument mit einem individuell wechselnden Seitendruck gehalten wird. Wer druckkräftig schreibt, bedarf eines grösseren Seitendrucks als wer eine Fadenschrift liefert. Der Seitendruck hängt nicht nur vom Druck gegen die Schreibfläche ab, sondern auch von der Versteifung der am Schreibakt beteiligten Muskeln des Armes. Die wissenschaftliche Graphologie hat angefangen, diesen Momenten des Schreibvorgangs ihre Beachtung zu schenken. Bei der Tätigkeit des Chirurgen und des Zahnarztes bemüht sich die Hand darum, ein optimales operatives Ziel zu erreichen, wobei weder zu weit gegangen noch zu wenig geleistet werden darf. Manche Zahnärzte sind wegen ihrer leichten Hand beliebt, offenbar ist bei ihnen die Relation zwischen den Sicherheitsmarginalien der verschiedenen an der Arbeit beteiligten Muskelgruppen besonders günstig.

Andere motorische Leistungen, auf die der Begriff des Sicherheitsmarginals anwendbar ist, sind die sportlichen. Man denke



an den Sprung über das Seil, an den Weitsprung mit festgelegtem Ziel, an Würfe mit dem Ball oder dem Speer. Das Prinzip des Sicherheitsmarginals ist aber nicht auf die Motorik beschränkt. Um welche seelische Leistungen es sich auch handle, immer muss ja bei der Erreichung eines Zieles — sei es mnemischer, sei es intellektueller Natur — eine untere Grenze nicht nur erreicht, sondern überschritten werden. Die klassische Gedächtnisforschung hat stets darauf hingewiesen, dass es die schwächste Stelle in einer zu lernenden Silbenreihe ist, die über die Lernarbeit entscheidet. Welche Bedeutung Erwägungen über das Sicherheitsmarginal für den Studenten besitzen, der vor einem Examen steht, ist offenbar. Der eine wagt sich mit einem Minimum von Kenntnissen in die Höhle des Examinators, der andere opfert unnötig viel Zeit und Kraft bei der Vorbereitung, weil er das Marginal zu hoch nimmt. Unnötig zu sagen, dass es sich hier vielfach um neurotische Sicherheitstendenzen im Sinne des Adlerschen Individualpsychologie handelt. Ein ausgezeichnete Psychologe, den ich persönlich gekannt habe, sollte sich im Zusammenhang mit dem Doktor-Examen in Psychologie auch für ein Examen in der Geschichte der Philosophie vorbereiten. In dieses letztere Examen konnte auch eventuell eine Frage nach dem Einfluss der arabischen Philosophie auf die westliche Philosophie eingehen, aber das war nicht einmal sicher, und hätte er sie nicht beantwortet, so hätte ihm das sicher nicht den Hals gebrochen. Wie bereitete sich dieser Psycholog auf das Examen vor? Unter anderem dadurch, dass er anfang, arabisch zu studieren. Unser Psychologe hat übrigens nie sein Examen abgelegt, sondern wurde Dr. honoris causa. Er war hochgradig Neurotiker, und wenn Goethe einmal gesagt hat, dass der typische Fall immer der allgemeine Fall ist, so gilt das auch hier. Der neurotische Charakter arbeitet in der Regel mit allzu hohem Sicherheitsmarginal und erschöpft dabei seine Kräfte auf unfruchtbare Weise.

Zu den Leistungen, die von störenden Faktoren aller Art bedroht sind, rechnen auch Vorlesungen und Vorträge. Es wäre verlockend, einmal eine Enquête darüber veranstalten, welche Massnahmen Vortragende überhaupt ergreifen, um ihren Erfolg technisch so weit als möglich im Voraus zu sichern. Ich kenne Kollegen, die zwar völlig frei vortragen und dies auch selbst bei exponierten Situationen zu tun wagen, die aber wenigstens

die letzten Sätze des Vortrags schriftlich niederlegen und vorlesen, weil sie damit rechnen, zum Schluss nicht mehr genügend Spannung zu besitzen, um die effektivste Formulierung für ein happy end zu finden. Bei einem Lichtbildervortrag kann man gar nicht genug tun, um sich vor der Tücke des Objekts zu schützen. Hier kommt es zuweilen zu okkulten Phänomenen, indem die Lichtbilder sich spontan auf den Kopf stellen. Ich habe mir oft die Frage vorgelegt, wie es zu verstehen sei, dass ein Redner so gut wie nie einen Vortrag abbrechen muss, weil sich zu unpassender Zeit gewisse körperliche Bedürfnisse einstellen. Ich habe so etwas nie erlebt, obwohl ich wohl Tausenden von Vorträgen beigewohnt habe.

Ich hatte in Aussicht gestellt, das Problem des Sicherheitsmarginals in den grösseren Zusammenhang des Sicherungsproblems des Menschen zu stellen. Die Sicherheit und Existenz des Menschen ist teils durch solche Umstände bedroht, die mit seiner Teilhabe am Leben unaufheblich verbunden sind, also durch körperliche und geistige Erkrankung, durch Altern, durch Unglücksfälle, bei Frauen auch durch das Risiko bei Geburten. Die Natur bedroht den Menschen mit Hungersnot infolge von Trockenheit und Ueberschwemmung, durch Erdbeben und Vulkanausbrüche. Von anderer Art sind diejenigen Gefahren, die mit den sozialen Verhältnissen zusammenhängen, Bedrohung mit Verarmung wegen Arbeitslosigkeit oder aus anderen Gründen, wie etwa Alkoholismus. Der Dichter Max Brod unterscheidet edle und unedle Not. Unedle Not ist diejenige Not, die durch menschliche Organisation beseitigt werden könnte, wenn nicht die Trägheit des Herzens bestünde. Dass man auf der ganzen Welt in dieser Hinsicht auf gutem Wege ist, wobei man die feierliche Erklärung über die Menschen rechte, die die UNO proklamiert hat, zu dem geeignetesten Wegweiser wählen kann, ist einer der wenigen Lichtpunkte unserer düsteren Zeit. Mit der schlimmsten und schändlichsten Not, mit dem Krieg, sind wir noch nicht zurechtgekommen.

Die Sicherungen, die der Mensch für seine Zukunft trifft, sind teils genereller teils individueller Natur. Sie haben teils direkten, teils indirekten Charakter. Direkt und individuell sind alle Massnahmen, die man persönlich ergreift, um sich für die Aufgaben, die einen vermutlich in der Zukunft erwarten, geschickt zu machen. Indirekter Natur sind z.B. alle Versiche-

rungen, die man eingeht, wie Unfall-, Krankheits-, Einbruchs-, Feuer- und Lebensversicherungen. Die Versicherungsgewohnheiten variieren beträchtlich von Land zu Land, abhängig davon, wie es mit der sozialen Fürsorge steht, die der Staat organisiert. Hier wären interessante den Nationalcharakter betreffende vergleichende Untersuchungen am Platz. In direkter Weise bereitet man sich auf zukünftige Aufgaben durch Erwerbung von Fertigkeiten und Kenntnissen vor; ich kan lesen, schreiben, rechnen, schwimmen, tanzen, ich kann mich in einigen fremden Sprachen ausdrücken und verständlich machen, ich habe einige Berufskentnisse. Je mehr ich kann, um so ruhiger kann ich der Zukunft entgegensehen.

Für das Problem, auf welche Weise wir uns für zukünftige Situationen zu sichern versuchen, sind Untersuchungen von beträchtlichem Interesse, die vor längerer Zeit mein Schüler Hans Keller über die Psychologie des Zukunftsbewusstseins durchgeführt hat <sup>5)</sup>. Das Zukunftsbewusstsein ist ja überhaupt gegenüber dem auf die Vergangenheit bezogenen Gedächtnis von der Psychologie stiefmütterlich behandelt worden. Wir machen uns in der Regel nicht klar, wie sehr unser Handeln dauernd durch unsere Vorstellung von der Zukunft mitbestimmt wird. Wenn ich die Gewissheit hätte — sagt Keller — dass mein Leben heute abend ein Ende fände, so würde dies mein gegenwärtiges Leben total umstürzen.

Können heisst: in allen kritischen Punkten der Zukunft den auftretenden Schwierigkeiten und Gefahren gewachsen zu sein. Erfahrung ist Wissen um die kritischen Zeitpunkte der Zukunft. Es kommt nach Keller zu Misserfolgen in zwei Füllen: 1. Wenn ich das Eintreten eines kritischen Zeitpunktes nicht bemerkte. 2. Wenn ich den Eintritt des kritischen Zeitpunktes zwar bemerke, aber versage. Beim Versagen kann man drei Fälle unterscheiden: : 1. Ich versage, weil ich nicht bei der Sache, abgelenkt, müde oder träge bin. 2. Ich versage in der Schicht der psychologischen Funktionen. Die assoziative Tätigkeit versagt, es kommt zu Täuschungen und Fehlleistungen der Auffassung und des Gedächtnisses, zu Verwechslungen und Hemmungen. 3. Ich versage in der Schicht des körperlichen Ichs,

---

<sup>5)</sup> Hans Keller, Psychologie des Zukunftsbewusstseins. Zeitschrift für Psychologie, 1932.



der Muskelkraft, des Tempos und der Reaktionszeit. Hier gilt es jede Art körperlich bedingter Schwäche.

Das Zukunftsbewusstsein ändert sich mit der jeweiligen kulturellen Lage, mit der Ausgeprägtheit des historischen Bewusstseins und in Abhängigkeit von noch vielen anderen Faktoren. Es ändert sich auch in der Krankheit. (H. Keller.) Der Depressive ist wie von der Zukunft abgeschnitten, er fühlt sich durch das Vergangene überwältigt und gebunden. Er lehnt Trost und Zuspruch des Arztes ab. Der psychopathisch Verstimmte fühlt sich nicht von der Zukunft abgeschnitten, sondern nur von den gegenwärtigen und zukünftigen Ereignissen, die er fürchtet, bedroht. Er wünscht Trost und Hilfe und nimmt sie gern vom Arzt an.

Während unsere Grundhaltung gegenüber der Vergangenheit das Wissen ist, steht alles Zukünftige grundsätzlich unter dem Zeichen der Unsicherheit und damit der Sorge. Um die Sorge zu bannen, um Zuversicht gegenüber den launischen Geschehnissen der Zukunft zu gewinnen hat die Phantasie des Menschen seit undenklichen Zeiten ideelle Gehäuse gebaut, um in ihnen Zuflucht zu suchen. In der erhabensten Form bemühen sich die grossen Weltreligionen um seine Tröstung. In der Philosophie hat dies Kant in der lautersten Form versucht in seiner berühmten Sicherheitsmaxime: „Die echte, mit der Religion allein vereinbare Sicherheitsmaxime (lautet): was als Mittel oder als Bedingung der Seligkeit mir nicht durch meine eigene Vernunft, sondern nur durch Offenbarung bekannt (ist), übrigens aber den reinen moralischen Grundsätzen nicht widerspricht, kann ich zwar nicht für gewiss glauben und beteuern, aber auch ebensowenig als gewiss falsch abweisen.... In dieser Maxime ist wahrhafte moralische Sicherheit, nämlich vor dem Gewissen und mehr kann von einem Menschen nicht verlangt werden.“<sup>6)</sup>

Heinrich IV. soll sich einmal bei einer Religionsdebatte des Prinzips der Sicherheitsmaxime auf eine Weise bedient haben, die nicht ganz auf Kantschem Niveau lag. Als von protestantischer Seite erklärt wurde, ein ehrlicher Katholik könne gerettet werden, während von der Gegenseite erklärt wurde, dass alle Protestanten zur Verurteilung bestimmt seien, äusserte Heinrich

---

<sup>6)</sup> Immanuel Kant, Die Religion innerhalb der Grenzen der reinen Vernunft. 4. Stück, 2. Teil, 4.



IV.: „Die Klugheit gebietet, dass ich deren Religion angehöre und nicht Eurer, denn wenn ich ihnen angehöre, so werde ich nach ihrer und Eurer Lehre gerettet, gehöre ich Euch an, so werde ich nach Euch gerettet, aber nicht nach ihrer.“

Meine bisherigen Ausführungen wollten und konnten kaum mehr als ein Forschungsprogramm sein. Es war persönlich gefärbt, insofern es phänomenologisch betont war und in den luftigen Höhen der Philosophie landete, aber ich bin es ihnen schuldig wenigstens anzudeuten, wie wir versucht haben mit Hilfe des Experimentes und in bescheidenerem Umfang von Enquêtemethoden dieses Forschungsprogramm zu verwirklichen. Ich beginne mit den Experimenten über gehobene Gewichte.

Um den Seitendruck zu bestimmen, den eine Versuchsperson auf den Handgriff eines Gewichtes ausübte, wurde ein einfacher Apparat konstruiert. Wir bauten für unseren Zweck eine jener einfachen Küchenwagen um, die mit Feder arbeiten und bei denen der auf die Wagschale ausgeübte Druck an einer Skala abgelesen wird. Die Versuchsperson umspannte zwei mit Haken versehenen Handgriffe, an denen die Gewichtsgefässe befestigt wurden. Je nach dem Druck, der auf die Handgriffe ausgeübt wurde, variierte deren gegenseitige Entfernung. Sie betrug bei dem kleinsten vorkommenden Gewicht von 1 kg etwa 8 cm und bei dem grössten von 10 kg etwa 4 cm, schwankte also um 4 cm. Die Schwankungsbreite ist nicht sehr gross, darum dürfte die in dieser Ungleichheit liegende Fehlerquelle nicht sehr bedenklich sein. Sie hätte sich vermeiden lassen bei Verwendung von piezoelektrischen Einrichtungen, die aber kostspielig sind. Der Druck, der die Versuchsperson auf die Handgriffe ausübte, konnte vom Versuchsleiter ohne weiteres an der Skala der Wage abgelesen werden. Natürlich war es der Versuchsperson nicht gestattet, selbst Kenntnis von den Druckwerten auf der Skala zu nehmen.

Gewichte von 1 bis 10 kg wurden in zufälliger Ordnung und an mehreren aufeinanderfolgenden Tagen gehoben. Es war der Versuchsperson also nicht möglich, sich auf eine steigende oder fallende Skala einzustellen, vielmehr musste sich die Hand in jedem einzelnen Fall dem individuellen Gewicht anpassen. Es ergab sich, dass der Seitendruck regelmässig mit dem gehobenen Gewicht zunahm, regelmässig in dem Sinne, dass jedem Wert ein grösserer folgte. Dies ist keineswegs selbstverständlich, da

der Seitendruck immer sehr wesentlich über dem minimalen lag. Nur in seltenen Fällen kamen Ausnahmen von dieser Regel vor.

Je kleiner die Beanspruchung der Hand bei der Gewichtshebung ist, um so mehr erfolgt die Anpassung des Seitendruckes automatisch. Je grösser das Gewicht, um so mehr greift der Wille in die Hebung ein. Das Maximum ist indessen bei 10 kg längst nicht erreicht.

Das Sicherheitsmarginal beträgt bei der Hebung eines Gewichts von 1 kg durchschnittlich etwa 92 % und sinkt proportional mit dem Gewicht bis auf 27 % bei einem Gewicht von 10 kg.

Die individuellen Unterschiede des Sicherheitsmarginals sind, wie schon früher angedeutet wurde, enorm gross. Während das Marginal für 1 kg bei einer Versuchsperson 97 % betrug und für 10 kg bei 50 % lag, betrug es bei einer anderen für 1 kg 6 % und für 10 kg 13 %. Die letztere Versuchsperson fiel durch die grosse Labilität ihres Verhaltens bei den Experimenten auf, eine Labilität, die auch ihr übriges Verhalten kennzeichnete. Ihr Sicherheitsmarginal lag, wie schon gesagt, für 1 kg bei 6 %, schwankte hin und her mit 54 % als dem Maximum und schloss bei 10 kg mit den bereits genannten 13 %. Eine besonders regelmässige Versuchsperson begann mit 96 % bei 1 kg und ging sehr regelmässig bis auf 38 % herab bei 10 kg. Es liegt nahe zu untersuchen, wie hoch die Sicherheitsmarginalen bei verschiedenen Leistungen miteinander korreliert sind. Es wäre sicher sehr schön, falls sich herausstellen sollte, dass man Tendenzen des Sicherheitsverhaltens mit Hilfe des Sicherheitsmarginals bei gehobenen Gewichten als Test bestimmen könnte.

Da es wahrscheinlich ist, dass manche Formen von Schreibkrampf bedingt oder wenigstens mitbedingt sind durch zu hohes Sicherheitsmarginal des Seitendruckes, so haben wir Experimente begonnen, bei denen dauernd mit genau festgelegtem Seitendruck geschrieben wurde. Der Federhalter besteht aus zwei gegeneinander beweglichen Teilen, auf die ein variabler Druck ausgeübt werden kann. Die Druckstärken sind geeicht, die Kontrolle wird in der Weise ausgeübt, dass bei einem vorher festgelegten Druck ein Strom den Federhalter passiert und eine Glühlampe zum Leuchten bringt. Die Versuchsperson hat die Aufgabe so zu schreiben, dass die Lampe dauernd leuchtet. Lässt man nun mit einem Seitendruck schreiben, der den ge-

wohnten der Versuchsperson wesentlich übersteigt, so stellen sich bald krampfartige Schmerzen ein, die ein Weiterschreiben fast unmöglich machen.

Um Experimente aus dem Gebiet des Schreibens von ganz anderer Art handelt es sich bei folgenden. Die Sicherheit des Schreibaktes, der zu einer gut lesbaren Schrift führt, hängt offenbar abgesehen von anderen Faktoren auch von dem Verhältnis ab, in dem die gesamte Schreibzeit eines Textes zu den Schreibpausen steht, die zwischen den einzelnen Wörtern oder innerhalb der Wörter gemacht werden. Dieses Verhältnis lässt sich mit Hilfe des Skriptochronographen bestimmen <sup>7)</sup>. Es gibt merkwürdigerweise Individuen, bei denen die Pausen nicht weniger als 64 % der gesamten Schreibzeit absorbieren, aber selbst der Durchschnitt für 35 Versuchspersonen lag bei einem Text von 50 Wörtern mit 209 Buchstaben immer noch bei 30 %, ein Wert, den man auf Grund der täglichen Erfahrung beim Schreiben sicher für zu hoch halten würde. Lässt man nun die Versuchspersonen schneller schreiben, d.h. fordert man Verkleinerung des Sicherheitsmarginals, so können die Versuchspersonen dieser Instruktion ohne grössere Schwierigkeit wenigstens für eine gewisse Zeit nachkommen, ohne dass die Qualität der Leistung darunter leidet. Hierbei nimmt die Schreibzeit um etwa 21 % ab, doch ist von besonderem Interesse, dass der Zeitgewinn fast ganz durch Verkürzung der Pausen erzielt wird, dass dagegen die Zeit für die eigentliche Schreibleistung nur um 8 % abnimmt. Das Gesagte gilt übrigens auch für Stenographie, was ja nicht nur psychologisch interessant ist, sondern auch zu wichtigen pädagogischen Konsequenzen führen kann. Ich halte es für wahrscheinlich, dass ähnliches wie für das Schreiben auch für das Sprechen gilt. Wir können wohl nur innerhalb recht enger Grenzen die Produktion von Wörtern beschleunigen. Wenn wir, wie das gelegentlich am Schluss eines Vortrages geschieht, schneller sprechen, so gelingt das in erster Linie dadurch, dass wir die Pause zwischen den einzelnen Wörtern verkürzen.

Zum Schluss komme ich auf einige Ergebnisse zu sprechen, die mit Hilfe der Enquêtemethode erhalten worden sind. Die Fragen bezogen sich auf Erwartungen, die die Befragten — alle

---

<sup>7)</sup> D. Katz, Der Skriptochronograph. Beiheft Nr 18 der Schweiz. Zeitschrift für Psychologie und ihre Anwendungen.



Studierende der Psychologie — hinsichtlich gewisser in der näheren oder fernerer Zukunft liegenden Vorkommnisse hegten.

Unser Leben gerät mehr und mehr unter die Herrschaft der Statistik. Man hat Sterblichkeitstabellen und Statistiken über die Häufigkeit, mit der Morde, Selbstmorde, Eheschliessungen, Diebstähle, Feuerbrünste, Unfälle, Krankheiten aller Art und andere Geschehnisse vorkommen. Statistisch gesehen haben wir in einem bestimmten Alter nach der Sterblichkeitstabellen Aussicht uns noch so und so vieler Jahre des Daseins zu erfreuen, haben wir mit dieser oder jener Wahrscheinlichkeit damit zu rechnen, einem Unfall, einer Krankheit, einem Diebstahl ausgesetzt zu werden, aber auch — um etwas erfreulicherer zu vermehren — das grosse Los zu gewinnen. In allen diesen Fällen handelt es sich ja um Wahrscheinlichkeiten höheren und niedrigeren Grades, die sich auf unsere Manipulation des Sicherheitsmarginals auswirken. Ich habe mir nun die Frage vorgelegt, wie sich eigentlich die statistisch geltenden Gesetze subjektiv und rein gefühlsmässig in unserem alltäglichen Verhalten und damit auch in unserem Verhalten und Planen für die Zukunft widerspiegeln.

An den verschiedenen Erhebungen nahmen 75—91 Personen teil. Eine erste Frage bezog sich auf das Resultat eines Examens, das unmittelbar auf die Befragung folgen sollte. Andere Fragen bezogen sich auf die Lebensaussichten bei einer Operation durchzukommen, auf das Umgehen mit Geld und auf das zeitliche Sicherheitsmarginal bei einer Reise mit einem Zug. Es wurde weiter die allgemeine Frage gestellt, ob man Pessimist oder Optimist oder keins von beiden sei und schliesslich ob man Sinn für die Devise habe „gefährlich zu leben“.

Hier folgen einige der Resultate, die wir bekommen haben. An einem schriftlichen Examen nahmen 91 Studenten teil, darunter 55 Männer und 36 Frauen. Unter ihnen waren 33 Optimisten — 15 Männer und 18 Frauen — 15 Pessimisten — 10 Männer und 5 Frauen — sowie 43 neutraler Weltanschauung, darunter 29 Männer und 14 Frauen. Es schien sich hier die auch sonst gemachte Erfahrung zu bestätigen, dass die Frauen mehr zum Optimismus zu neigen scheinen. Es wurden nun alle Teilnehmer am Examen gebeten, schriftlich und in verschlossenem Kouvert sich darüber zu äussern, ob sie erwarteten, das Examen zu bestehen oder nicht zu bestehen oder aber — was eine Spezial-



tät der schwedischen Examenseinrichtung ist — einen sogenannten „Rest“ zu bekommen. Wer in einem Examen einen Rest bekommt, muss nur einen Teil desselben wiederholen. Um die Beantwortung der Fragen soweit als möglich gegen entstellende Einflüsse zu sichern, wurden die Antworten von einem Vertrauensmann der Studenten eingesammelt und dieser wurde „coram publico“ dahin instruiert, dem Examinator die Antworten erst nach Abschluss der Prüfung der Arbeiten zu übergeben. Erfahrungsmässig bestehen etwa 70 % die Prüfungen, etwa 10 % sind Versager und die verbleibenden 20 % bekommen einen Rest. Dies wurde den Prüflingen bekannt gegeben, und sie sollten sich darüber äussern, in welche Kategorie sie wohl geraten würden. Es zeigte sich, dass 34 glaubten das Examen zu bestehen, 12 nicht zu bestehen und die übrigen 45 einen Rest zu bekommen, ausgedrückt in Prozent sind das etwa 37 %, resp. 13 % und 50 %. Es ist also eine sehr starke Verschiebung im Sinne weniger günstiger Resultate erwartet worden. Von den vier besten war nur einer Optimist, zwei waren neutral und einer sogar Pessimist. Ein Pessimist glaubte nicht bestehen zu können, erhielt aber ein sehr gutes Resultat. Man hat guten Grund anzunehmen, dass das Resultat dieses Versuchs nicht durch eine Rücksicht auf den Examinator beeinflusst worden ist, und dass in ihm ein Verhalten halb abergläubischer Art zum Ausdruck kommt. Man möchte das Schicksal nicht herausfordern. Es ist dieselbe Haltung, die uns im täglichen Leben in gewissen Situationen „unberufen“ sagen oder doch denken lässt. So viel über einen tastenden Versuch, etwas über die Einstellung zur unmittelbar bevorstehenden Zukunft zu erfahren. — Ganz im Gegensatz dazu stand die nächste Frage, die darauf ausging zu erkunden, welche Lebensaussichten die Studenten zu haben glaubten. Um Schockwirkungen zu vermeiden, die die Frage: „Wie alt glauben sie werden?“ zur Folge haben mochte, habe ich ihr eine solche Form gegeben, die zwar nur zu weniger exakten Resultaten führen konnte, aber doch leichter akzeptiert wurde. Sie lautete: Glauben Sie das Jahr 2000 zu erleben? Nur bei einem der 61 Befragten hat auch diese Frage noch irritierend gewirkt. Er antwortete: Hieran habe ich nie gedacht und will auch nicht daran denken. Von 61 haben 31 die Frage mit ja beantwortet, 22 mit nein, 8 haben nicht geantwortet. Von den Optimisten haben 50 % mit ja geantwortet. Bei einem mittleren Alter von 28 Jahren be-

deutete dies eine Lebenserwartung von 78 Jahren. Von den Pessimisten haben nur 33 % mit ja geantwortet, und doch betrug ihr Durchschnittsalter bei der Befragung nur 25 Jahre. Diese Befunde sind nicht ganz ohne praktische Bedeutung für die Arbeit der Lebensversicherungen, nach deren Erfahrung bei den Versicherungsnehmern eine deutliche Tendenz vorhanden ist, sich nicht mit dem Gedanken an den Tod zu beschäftigen oder dessen Risiko zu unterschätzen.

Eine Frage galt der Einschätzung des Risikos einer Operation. Sie lautete: Bei einer ärztlichen Untersuchung sagt Ihnen der Arzt, dass Sie sich einer Blinddarmoperation unterziehen müssen. Für wie gross halten Sie die Sterblichkeit ausgedrückt in %. Die Antwort war im Durchschnitt 2 %, dabei lag sie für die Optimisten bei 1.8 % und für die Pessimisten bei 3.3 %. Bei einer anderen Gruppe wurde die Frage dahin modifiziert, dass gefragt wurde: ein naher Verwandter erfährt bei einer ärztlichen Untersuchung, dass er sich einer Blinddarmoperation unterziehen muss. Für wie gross halten Sie die Sterblichkeit? Der Wert lag im Durchschnitt bei 1.9 %, ist also etwas niedriger, als wenn es die eigene Person gilt.

Die Studenten wurden gefragt, ob sie über ihre Einnahmen und Ausgaben Buch führten. 21 % antworteten mit ja, 68 % mit nein, 11 % tun dies nur gelegentlich. Es müsste für die Attitüden zu Geldangelegenheiten sehr aufschlussreich werden, ähnliche Untersuchungen in anderen Ländern durchzuführen. Der Schweizer Psychiater E. Bleuler, der 1911 gestorben ist, soll einmal gesagt haben, es sei belastend, wenn sich in der Anamnese eines Schweizer Patienten herausstellte, dass er kein Sparkonto habe, es sei aber eine Anomalie, wenn ein russischer Patient ein solches habe.

Ein Moment der Befragung betraf das Sicherheitsmarginal bei der Benutzung der Eisenbahn. Die Frage war so formuliert: Sie haben die Absicht ohne Bagage mit dem Zug, der Stockholm um 13 Uhr verlässt, nach Malmö zu fahren. Wann finden Sie sich bei dem Bahnhof ein, wenn Sie kein Platzbillet haben? Eine andere Gruppe wurde vor dieselbe Frage gestellt, aber mit dem Unterschied, dass ein wesentlich kürzeres Reiseziel mit demselben Zug gesteckt wurde und zwar Södertälje. Es stand also mehr auf dem Spiel bei der Fahrt nach Malmö. Tatsächlich ergab sich, dass sich die Reisenden nach Malmö durchschnittlich

21 Minuten vor Abgang des Zuges einfanden, die Reisenden nach Södertälje dagegen erst 13 Minuten vor Abgang. Die Pessimisten fanden sich sogar 25 Minuten vor Abgang des Zuges nach Malmö ein.

Eine letzte Frage, die an die Studenten gestellt wurde, war, ob sie Sinn für die Devise hätten, gefährlich zu leben und damit verbunden war die Frage, ob sie einmal einen Unfall gehabt hätten. Von den Männern haben 39 % einmal einen Unfall gehabt und 19 % lieben es gefährlich zu leben. Von den Frauen haben etwas mehr einen Unfall gehabt, nämlich 42 %, aber nicht weniger als 37 % lieben es gefährlich zu leben.

So viel über die Bemühungen, die der Durchführung des Forschungsprogrammes über das Sicherheitsmarginal und das allgemeine Sicherungsproblem gewidmet worden sind, doch ist der Bericht keineswegs erschöpfend. Unter anderm fehlt dabei eine Untersuchung, die im Auftrage der Unesco durchgeführt worden ist unter dem Titel: Coordinating the International Organizations in the Interest of the Unesco Project „Tensions affecting International Understanding“, und die ein konstruktives Program darüber enthält, was die internationalen Organisationen zur Beseitigung von internationalen Spannungen und damit zur Milderung der Kriegsgefahr tun könnten. Den internationalen Gruppen, die der Wissenschaft und Forschung dienen, fällt hierbei eine besonders wichtige Rolle zu. Die Adepten einer Wissenschaft sind ja auf besonders enge Weise miteinander verbunden. Sie haben häufig dieselben Bücher gelesen, sie sprechen dieselbe wissenschaftliche Sprache, ihr Hauptinteresse haben sie gemeinsam, sie haben häufig dieselben Freunde, seien es nun lebende oder die grossen Gestalten aus der Geschichte der Wissenschaft. So sind auch wir Psychologen eine grosse internationale Bruderschaft. Und solcher wissenschaftlichen Bruderschaften gibt es in der wissenschaftlicher Welt erstaunlich viele. Es sollte sich ein Versuch lohnen diese im Interesse der Sicherung des Friedens zu organisieren, eventuell auch über die Köpfe der Regierungen hinweg.

Wenn wir im nationalen Raum die Polizei, die Feuerwehr, den Wetterdienst und andere Einrichtungen organisieren, so schaffen wir Sicherungen gegen das Verbrechen, gegen die Feuersgefahr, gegen Wetterschäden und andere Risiken, Sicherungen die dem Umfang nach völlig beherrschbar und rational



anpassbar sind. In der internationalen Politik versucht man sich durch die Organisation militärischer Kräfte ein möglichst grosses Sicherheitsmarginal zu verschaffen, aber jede Vergrösserung dieses Marginals im nationalen Raum hat unheimliche internationale Folgen. Wie der individuelle Neurotiker bei seinem Streben nach einem zu hohen Sicherheitsmarginal daran gehindert wird, sich des Lebens zu erfreuen, so kann schliesslich infolge der sich gegenseitig steigernden Rüstungen die Welt von globaler Stagnation und Lähmung befallen werden, oder es kommt, ganz wie beim individuellen Neurotiker, zu einer Katastrophenreaktion.

#### ZUSAMMENFASSUNG

Der Ausdruck Sicherheitsmarginal oder Sicherheitsfaktor ist der Technik entnommen. Die Übertragung dieses Begriffes ins psychologische hat sich als sehr fruchtbar erwiesen. Um nur ein Beispiel anzuführen, sei auf diejenigen Personen hingewiesen, die bei bestimmten Gelegenheiten zu früh kommen, die zu spät kommen und die rechtzeitig kommen. Die ersten machen sich und evtl. einer Organisation durch ihr vorzeitiges Kommen Schwierigkeiten, ihr Sicherheitsmarginal ist zu gross. Die zweiten versagen durch ihr zu spät kommen, ihr Sicherheitsmarginal ist zu gering. Schliesslich, die dritten haben ein angemessenes Sicherheitsmarginal. Das Verhalten der drei Individuen ist nicht durch einen Zufall bedingt.

Das Sicherheitsmarginal ist der Gegenwart zugewandt, dem gegenwärtigen Raum und der gegenwärtigen Zeit, im Gegensatz zu der rückwärts gerichteten Arbeit des Gedächtnisses. Das Problem des Sicherheitsmarginals gehört in erster Linie in die allgemeine Psychologie, doch ist es auch besonders differentialpsychologisch anwendbar.

Es ist aber nicht nur die Psychologie, in der das Sicherheitsmarginal Anwendung findet, es ist auf jedem Gebiet der menschlichen Beziehungen brauchbar und kann als Test angewandt werden.

#### SUMMARY

The term: "safety margin" or "safety factor" has been borrowed from the technical field. The transfer of this concep-



tion to the psychological field has proved to serve a variety of purposes. To give an example one might refer to those people who at certain occasions will come too early. Others will be late, others again will be in time. The former group will cause trouble for themselves and possibly for some organisation — their safety margin is too large. The second group are at fault because they arrive late — their safety margin is too small. The third group finally have a correct safety margin. The way of acting of the three individuals is not accidental. The safety margin is based on (connected with) the present — the present space and the present time, contrary to the reverse action of the memory.

The problem of the safety margin belongs in the first instance to general psychology, but is also particularly suited to differential psychology.

It is however not only in psychology that the safety margin finds application. It can be used in every field of human relationship and can be applied as a test.

# AN EXPLORATION OF THE INFLUENCE OF PERSONAL RELEVANCE UPON STATEMENTS OF AESTHETIC PREFERENCE

BY

WILLIAM BEVAN JR. AND GRETHA SEELAND

*Emory University and the University of Oslo*<sup>1)</sup>

## INTRODUCTION

If one were obliged to characterize the "Zeitgeist" of present-day experimental psychology with a single word or phrase, a rather convincing argument might be advanced to justify the label, "central-oriented". Learning theorists, in recent years, have been concerned less with the problem of operationally specifying external stimulus and response variables, taken up more with identifying the nature of intra-organismic middle states (e.g. 12). Physiological psychologists have been occupied less frequently by analyses of muscle twitches, more often by considerations of brain dynamics (e.g. 6, 9). Students of the motivation construct have been inclined less to direct their attention to the behavioral reflections of glandular functioning, more to center their thinking about the process of social conditioning (e.g. 8). Finally, workers in the area of perception have been interested less in determining the role of physical properties of the stimulus object in what and how things are seen, more in exploring the importance of the experimental-motivational complex in the structuring of the perceptual response (cf. 3, 4).

If this shift in emphasis from peripheral to central dynamics be taken as progress, then the special field of *experimental aesthetics* is marked by a striking lag in development. A glance through the chapter on aesthetic judgement in Woodworth's 1938 textbook of experimental psychology (14) suggests that experimenters up to that time were involved almost exclusively in attempts to establish correlations between the stimulus,

---

<sup>1)</sup> The present paper describes a study conducted in the fall of 1952 while the first writer was a research scholar at the Psychological Institute of the University of Oslo under the auspices of the Fulbright program.

physically defined, and the aesthetic reaction. A search through the *Psychological Abstracts* for the subsequent fourteen years reveals no noticeable change in trend. The comparatively few experimental papers summarized in the section on aesthetics deal predominantly with analyses of the physical stimulus and with devising measures of art appreciation and aptitude.

Why this lag in *experimental* aesthetics? We suspect that the answer is to be found in a noncritical acceptance of the assumption, traceable at least in certain aspects back to the Pythagoreans, that the capacity to arouse the aesthetic response is inherent in the form of the beautiful object. Any scientist in approaching his subject matter must ask himself at least two questions: What are the conditions that determine the occurrence of this phenomenon? How is it that these conditions achieve the effect? If the assumption of inherent beauty be held valid, the experimental aesthetician need only survey the several physical characteristics (in the case of the visual arts, e.g. hue, area, proportion, surface texture, figure-ground contrast, etc.) to have answers to both questions. While the validity of the assumption has *not* been universally accepted (cf. the concept of empathy applied to the problem of art appreciation (7), the emphasis on cultural standards (14, 391), a formulation in terms of problem solving (13), the approach through psychoanalysis (10) (11), the alternative views, often imaginative and stimulating, generally lack not only the prestige of age but, more importantly, empirical foundation sufficient to replace it. Indeed, it is curious to note the persistence with which the assumption has been retained by those who have accepted it. Although, for example, studies of color preference dating back to the Jastrow study of 1893 (14, 283) are striking in their lack of agreement, von Allesch (1) in 1925 still approached the problem by considering all discrepancies to be due to crudities in experimentation.

If beauty be inherent in the form of the beautiful object, then the psychological dynamics of the aesthetic response are wholly unique, for the vast body of experimentally derived knowledge in psychology lends credence to the view that the capacity of objects to evoke some particular sort of response grows out of an intercourse between the object and the responder. There is no convincing proof for such uniqueness. Nor is there much direct *experimental* evidence against it, although one may

suggest, of course, the anthropological reports of inter-cultural and inter-period differences in aesthetic preference as well as the lack of agreement within even relatively homogeneous sub-groups.

It is the view of the present writers that not only is the "inherent beauty" assumption invalid but also that it has placed unnecessary limitations on the experimental study of the aesthetic response. The aim of the present study was to support this view by experimentally demonstrating that statements of aesthetic preference may be influenced by a variable not classifiable as a "natural" property of the stimulus. More specifically stated, the present investigation was directed toward testing the assumption that aesthetic preference varies as a function of the personal relevance of the viewed object for the viewer <sup>2</sup>).

Accordingly, subjects were asked to rate relatively meaningless figures for aesthetic preference. Certain of the figures were then assigned meanings of generally high positive value, others meanings of an extremely negative sort, still others meanings of neutral or negligible affective tone. After each subject had thoroughly learned to identify each figure by its assigned meaning, he was again asked to rate the figures for aesthetic preference. Finally, pre-learning and post-learning ratings were compared. From the general assumption stated above the following more specific hypotheses were generated:

1. If aesthetic preference depends on the personal relevance (meaning) of the figure and if the figures are generally of low personal relevance (meaningless), then the mean of the initial ratings should not vary significantly from the indifference score.
2. If aesthetic preference changes with a change in personal relevance, then the average of the second ratings for those cards assigned positively-valenced meanings should show a significant shift toward the positive end of the scale; by the same token, the average of the second ratings of those cards assigned negatively-valenced meanings should shift toward the negative end of the rating scale.

---

<sup>2</sup>) The term *personal relevance* is substituted for *meaningfulness*, in order to emphasize the contention that the concept of meaning has motivational as well as cognitive implications.



3. If aesthetic preference changes with personal relevance, and if an increase in personal relevance implies an increase in motivation, then the second ratings of those figures assigned negative and positive meanings should be made with significantly greater precision than should those conveying relatively little meaning.

#### THE EXPERIMENTAL SITUATION

*Subjects:* Altogether 65 persons served as subjects. Fifty, half males and half females, were utilized in the preliminary task of selecting the positive and negative meanings to be assigned to a portion of the stimulus figures. The remaining fifteen, 7 males and 8 females, provided the ratings in the experiment proper. All had passed the *examen* for matriculation at the University of Oslo. Some were students, some office workers, others employed in skilled or semi-skilled work (a few of the female subjects were housewives).

*The stimulus figures:* The figures chosen for use were Chinese characters. There were selected (a) since it was assumed that they would have little or no meaning for the present subjects, and (b) since they have pictorial quality and at the same time, the capacity, because of their association with the language function, for readily assimilating rather specific meaning. Fifty different characters similar in complexity (number of strokes) and over all proportion were used. Each was drawn in black India ink (stroke width,  $\frac{1}{16}$ " ) on a separate sheet of heavy white drawing paper  $8'' \times 11\frac{1}{2}''$ . All figures were  $2\frac{3}{4}''$  wide and  $3\frac{3}{4}''$  high bordered by a rectangle  $5\frac{3}{4}''$  by  $6\frac{3}{4}''$ .

The stimulus cards were divided randomly into three groups: one of 20 cards, 3 of 10 cards each. The first was designated as a practice series to be presented on the first twenty trials of each test session in order to familiarize the subject with the rating procedure. The remaining three groups constituted the experimental series, one eventually being assigned positively valenced meanings, one negative meanings, and the third affectively neutral meanings.

*Selection of personally-relevant labels:* The meanings assigned after the initial rating were obtained in the following fashion: With the aid of a third person, the writers compiled a list of several hundred words and phrases believed to communicate

ideas of either highly positive or highly negative value. These were then carefully culled until there remained 116 items to which, it was confidently felt, the average subject would react. Fifty subjects were then asked to evaluate each of these items on a five point scale, ranging from 1: very unpleasant and repulsive, through the indifferent point at 3, to 5: very pleasant and attractive. Frequency counts were made of the responses to each item and the 10 items most frequently rated either 1 or 5 were picked as labels for the "positive" and "negative" stimulus figures (cf. Table I). The Chinese language includes a number

TABLE I  
Meanings assigned to figures of the experimental series

Positive	Neutral	Negative
1. Easter sunshine on the mountain meadows	1. beh	1. the gas chamber
2. world peace	2. mah	2. the raping of children
3. sun-warmed cliffs and salt sea	3. neh	3. the concentration camp
4. odor of heather on a warm day	4. lah	4. torturer
5. being in love	5. bah	5. fratricide
6. a summer holiday	6. geh	6. world war
7. good health	7. peh	7. a corpse full of crawling worms
8. passing an examination with the highest grade	8. ah	8. gestapo
9. coziness about a fire place	9. leh	9. the execution of a person
10. honesty	10. pah	10. a leprous finger in one's food

of sounds that have no literal meaning, instead convey strong feeling—e.g. surprise, delight, anger—or indicate commands or questions. Ten such sounds were applied to figures in the "neutral" group.

*Experimental sessions:* Testing was conducted during two approximately hour-long sessions separated by an interval of from several days to two weeks (average one week). Fifteen subjects were tested individually.

In the first session, the initial ratings of the figures were made.

The practice series was presented on the first twenty trials and was followed without interruption by a random arrangement of the experimental figures. When rating had been completed, the 30 experimental cards were separated from the practice cards and the ostensibly correct meaning of each explained to the subject. The subject was then given a chart containing these figures accompanied by their "meanings" and requested to learn to identify each.

On the second session, the cards were reviewed until the subject could successfully identify every member of the series. They were then recombined with the practice series and ratings of preference made a second time.

*Testing procedure:* After the subject had been made at ease in the testing situation he was given the following instructions:

"In this experiment we are interested in studying the aesthetic experience. Of course, the aesthetic experience itself is your own private affair and we cannot examine it directly. But we can deal with what you tell us about your reactions. We do know that people, when they are confronted by a "beautiful" thing, may say that it makes them feel good, that they are filled with warmth, that they stand in awe of it, etc.—in short, they say they have a pleasurable experience. I shall presently show you a series of Chinese characters. Such characters are frequently displayed as decorative pieces or works of art in Chinese homes. I want you to look at each, then try to answer the question, "What sort of feeling is aroused in me by this object?" Next, describe your reaction by giving me a number from the following scale:

1. The figure is *very very* pleasant and attractive.
2. *Very* pleasant and attractive.
3. Pleasant and attractive.
4. Slightly pleasant and attractive.
5. It arouses no reaction in me.
6. Slightly unpleasant and repulsive.
7. Unpleasant and repulsive.
8. *Very* unpleasant and repulsive.
9. *Very very* unpleasant and repulsive.

Try to give your *first* impression and work fairly rapidly."

After the initial rating had been completed, the subject received these additional instructions:

"Now I am going to show you some of the characters you have just seen and tell you what they mean. After I have given you the meaning several times, I want you to try to tell me what each is. Then when you leave today I would like you to take the following chart with you so that you may learn to promptly identify each character. Some of the characters have no literal meaning, that is, they do not refer to objects or situations, but are used instead to indicate emphasis, commands, questions, etc."

At the beginning of the second session the subject was informed that the cards which he had learned would be reviewed until he could correctly identify each. When he had met the criterion for learning, additional ratings of the sort made the first day were required of him.

#### RESULT AND DISCUSSION

*Initial ratings:* The major results are summarized in figure 1. Inspection indicates, first of all, that the ratings of aesthetic preference obtained during the first session are, on the average, of an indifferent sort. The average for all figures of the experimental series (4.84) does not differ significantly from the indifference point ( $5.00 - 4.84 = .16$ ;  $t = 1.88$ ;  $df = 29$ ,  $P .05$ ). If the assumption that the figures generally had no meaning for the subjects be valid, then these results may be taken as verification of hypothesis I. The absence of a shift in level of preference after learning in the group (neutral) with relatively nonsensical labels provides further substantiation of the proposition (cf. insignificant between-ratings  $F$ , Table II a.).

Not all figures, however, evoked initially neutral judgements from the subjects. Some were rated consistently above or below the neutral point (significant between-figures  $F$  ratios, Table II). If our general assumption concerning the relationship of meaning to preference be accepted, then one might expect such figures to possess some degree of meaning for the subject prior to that acquired through the training sequence. Spontaneous identifications (boxes, lanterns, knives, windows, etc.) occasionally made by the subjects during rating suggests such to be the case.

The presence of meaning in the experimental series on the first rating raises a certain doubt concerning the results to be described below. If our technique of randomization in the formation of the three sub-groups were inadequate, the relationship



of learning to the differences in rating subsequently noted might be held suspect. That this is unjustified, however, is revealed by a lack of between-groups difference on the first ratings: Between-groups  $F = .04$ ,  $df = 2$ ,  $P > .05$ .

TABLE II

Summaries of triple-classification analyses of variance performed on the initial and post-learning ratings of characters assigned neutral, positively- and negatively-toned meanings

A. Neutral Figures				
Source	df	Mean Square	F	P
Between subjects	14	4.38	2.37	< .01
Between ratings	9	7.49	4.05	< .01
Between figures	1	1.33	.72	> .05
Subjects by figures	126	2.77	1.50	< .05
Subjects by ratings	14	5.35	2.89	< .01
Figures by ratings	9	2.75	1.51	> .05
Residual	126	1.85		
B. Positive Figures				
Source	df	Mean Square	F	P
Between subjects	14	16.45	11.92	< .01
Between figures	9	6.19	4.49	< .01
Between ratings	1	140.08	101.50	< .01
Subjects by figures	126	2.35	1.70	< .01
Subjects by ratings	14	9.49	6.88	< .01
Figures by ratings	9	.31	.22	> .05
Residual	126	1.38		
C. Negative Figures				
Source	df	Mean Square	F	P
Between subjects	14	12.42	7.10	< .01
Between figures	9	3.60	2.06	< .05
Between ratings	1	215.05	122.89	< .01
Subjects by figures	126	2.90	1.66	< .01
Subjects by ratings	14	17.19	9.82	< .01
Figures by ratings	9	1.99	1.14	> .05
Residual	126	1.75		

*Shift in ratings with enhancement of personal relevance.* Inspection of Figure 1 also indicates that "injection" of meaning alters the average preferential rating. The reliability of this effect is revealed in Table II by the insignificant between-ratings

F ratio for the "neutral" group, the significant between-ratings F's for the "positive" and "negative" figures. The direction of shift is seen, furthermore, to parallel the direction of increased relevance; that is, figures which came to connote positive values for the subject increased in attractiveness; those associated in the between-ratings interval with values of a negative sort were judged less attractive. Hypothesis 2 thus appears also to be confirmed.

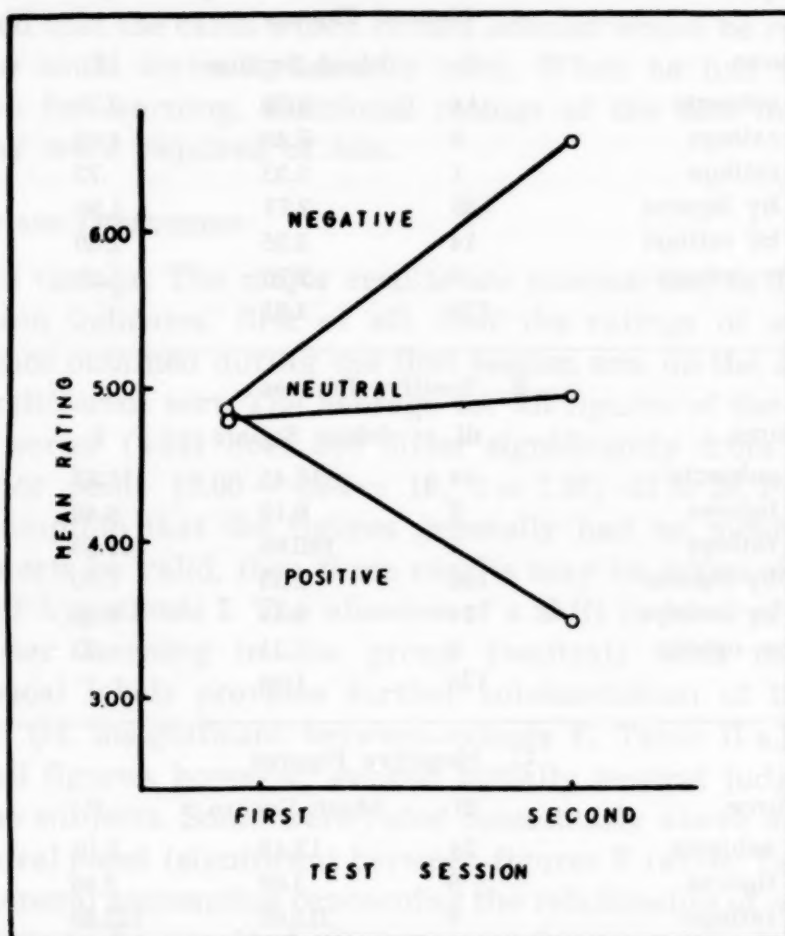


Fig. 1. Mean preferential ratings for each sub-group of the experimental series before and after learning the "meaning" of the figures

Figure 1 finally suggests the possibility that negative enhancement was more potent than positive in effecting the shift in preference. This coincides with expectation, since a much greater proportion of the judges used in the selection of the "meanings" agreed upon what was extremely unpleasant than

on what was extremely pleasant (mean frequency scores for 10 "negative" labels = 41.2; mean frequency scores for 10 "positive" labels = 29.6). Further analysis has shown, however, that this last conclusion is not warranted (1.69 (mean negative shift) — 1.37 (mean positive shift) = .32,  $t = .21$ ,  $df = 14$ ,  $P > .05$ ).

The importance of personal relevance for the aesthetic response is further highlighted by the insignificant figures-by-ratings source of variance in all of the analyses summarized in Table II. These statistics suggest that the figures themselves—that is, the physical configurations—exerted no detectable systematic influence upon the obtained shifts in preference.

*Precision in rating:* Hypothesis 3 states that enhancement of personal relevance is accompanied by increased precision in rating. This was tested by evaluating the reliability of the differences between the magnitudes of the standard deviations—the smaller the standard deviation the greater the precision—computed for each subject's first and second ratings of the several sub-groups in the experimental series. In none of these comparisons did the  $t$  ratio indicate the difference to be significant (neutral figures, difference between precision of first and second ratings expressed in S. D. units = .13,  $t = .75$ ,  $df = 14$ ,  $P > .05$ ; positive figures, difference = .24,  $t = 1.71$ ,  $df = 14$ ,  $p > .05$ ; negative figures, difference = .20,  $t = .92$ ,  $df = 14$ ,  $p .05$ ). Hypothesis 3 thus lacks substantiation.

Negative results demand some rationalization. The explanation that most readily suggests itself is that precision of judgement is not affected by personal relevance. This conclusion is weakened, however, by data that support, at least for other types of judgement, the precision hypothesis (2, 5). A failure to induce shifts in relevance of a magnitude adequate to effect measurable changes in precision provides another reasonable interpretation. Certainly the changes in preference obtained, while significant, probably represent only an extremely small proportion of the total dimension.

*Inter-subject differences in rating:* The between-subjects source in every analysis summarized in Table II is clearly significant. Inspection of the records of individual subjects revealed that, while the mean ratings of a large number on the first session fell near the indifference point, some tended to rate consistently above it, others below it. The subjects-by-figures

P		N	+	-
以 木巴	掬	慰	猱	椽
纂	糴	眾	追	喪
拒	愴	癰	磬	捶
闕	攢	楓	創	輻
憊	耨	權	侷	駿
確	霽	郡	窗	掘
苴	僂	矐	猓	阿
俱	輶	爽	裙	瘡
重	撞	椿	鵲	盟
銃	禡	蛾	懼	窘

Fig. 2. 'P' are the practice series, 'N' the "neutral" figures, + and - the "positive" and "negative" meanings



and the subjects-by-ratings interactions provide further interesting information. The former is significant in all analyses and indicates the importance of individual experience for preference level.

Should preferences derive from the "natural" properties of the figure—its color, its shape, etc., one might expect the response of all subjects to be relatively uniform. But the subjects-by-ratings interaction is significant for both the positive and negative groups, indicating instead that an increment in personal relevance does not affect the aesthetic preference of all subjects in the same fashion. Inspection of individual records, furthermore, showed the aesthetic preference of some subjects to be extremely sensitive to a shift in negative valence, while relatively insensitive to a positive change; in other cases, the opposite tendency was observed. In still others, preference was influenced to a marked extent by a shift in relevance of either a positive or negative sort, and in a fourth group, preference was affected hardly at all. These data suggest the probable fruitfulness of exploring further the correlation between personality variables—particularly, dominant styles or themas—and the nature of aesthetic judgements.

*General comments:* Several matters warrant further consideration in this report: The first concerns the validity of the assumption that our dependent variable is, *sensu stricto*, an aesthetic response; the second, the significance of meaning as a determinant of this response.

It may be contended by the students of art that the response we measure lacks the intensity, the complexity of content, perhaps the ambivalence, of the true aesthetic reaction. The only defense to such a position is that, in the final analysis, classification of such responses is arbitrary, depending on criteria derived from overt behavior. The rating scale of the present study follows precedent in its phraseology and is constructed to reflect affective tone in the judgements to which it is applied.

With regard to the second matter, the results clearly demonstrate that personal relevance (meaningfulness) influences what we have chosen to regard as an aesthetic judgement. They at the same time demand further clarification concerning whether on the one hand personal relevance is, along with the physical characteristics of the *objet d'art*, one of several classes of inde-

pendent variable, or, on the other, the *basic* determinant of preference (that is, object form and other variables are effective to the extent that they are meaningful to the responder). While confident answer requires further experimentation, all of the present data suggest the latter alternative: The generally indifferent initial ratings, the apparent meaningfulness of the figures initially rated consistently above or below the indifference point, the failure for the figures-by-ratings interactions to reach significance, the consistently significant between-subjects sources, subjects-by-figures and subjects-by-ratings interactions.

The present investigation is simple in its conceptualization; its results are in line with common expectation. If it has value, such is the case, perhaps, because it demonstrates the feasibility of applying the experimental method to an important area of study. It also invites future work on a wide variety of problems: e.g. What are the limits of personal relevance within which the relationship to aesthetic preference holds? How is the effect distributed within these limits? What is the relation between the ease with which an object becomes relevant and its capacity for evoking the aesthetic response? What aspects of objects are the most effective carriers of relevance—color, form, etc.? What determines precision in aesthetic judgement? What is the relative effectiveness of the several biologically-based and socially-induced values for evoking the aesthetic response? What is the relationship of degree of ambivalence to aesthetic preference? What is the relationship between other value systems and the subject's system of aesthetic value? How do personality schemas relate to aesthetic preference? Not only is experimental work on the relation of personality schemas to aesthetic preference feasible, it also provides a potentially fruitful medium for the study of personality structure itself.

#### SUMMARY

The present paper reports an experimental exploration of the relationship of personal relevance (meaningfulness) to aesthetic preference. After fifteen subjects had rated Chinese characters on a scale of attractiveness, they were told that certain figures represented positively valenced concepts, others negatively valenced concepts, still others had no literal meaning, were used only to indicate mood. When the meaning of each figure had

been thoroughly learned, ratings were again made. Analyses of the two sets of ratings provided the following conclusions.

1. When the figures were relatively meaningless, the response evoked was generally one of indifference.
2. When positive meanings were induced, attractiveness generally increased; when negative meanings were induced, the opposite shift was observed.
3. Induction of meaning had no effect on precision of rating.
4. Wide individual differences in rating pattern suggest the rich influence of personal experience upon aesthetic preference.

The implications of these findings for further work in this area are briefly commented upon.

#### REFERENCES

1. Allesch, G. J. von, *Die asthetische Erscheinungsweise der Farben*. Psychol. Forsch., 1925, 6, 1—91; 215—381.
2. Bevan, W. and W. F. Dukes, Value and the Weber constant in the perception of distance. Amer. J. Psychol., 1951, 64, 580—584.
3. Blake, R. R. and G. V. Ramsay, *Perception, an approach to personality*. New York: Ronald, 1951.
4. Cantril, H., *The why of man's experience*. New York: MacMillan, 1950.
5. Dukes, W. F. and W. Bevan, Accentuation and response variability in the perception of personally relevant objects. J. Personality, 1952, 20, 457—465.
6. Hebb, D. O., *Organization of behavior*. New York: Wiley, 1949.
7. Lipps, T., *Optische Streitfragen*. Z. Psychol., 1892, 3, 493—504.
8. Murphy, G., *Personality*. New York: Harpers, 1947.
9. Pitts, W. and W. S. McCulloch, How do we know universals? Bull. Math. Biophys., 1947, 9, 127—147.
10. Read, H., Psychoanalysis and the problem of aesthetic value. Int. J. Psychoanal., 1951, 32, 73—82.
11. Sachs, H., *The creative unconscious; studies in the psychoanalysis of art* (2nd Ed.). Cambridge, Mass.: Sci-Art. Publishers, 1951.
12. Tolman, E. C., Discussion. In Bruner, J. and D. Krech, *Perception and personality*. Durham, N. C.: Duke University Press, 1950, 48—50.
13. Weber, C. O., The aesthetics of rectangles and theories of affection. J. appl. Psychol., 1931, 15, 310—318.
14. Woodworth, R. S., *Experimental psychology*. New York: Holt, 1938, Ch. 16.

*From the Phonetics Laboratory of Amsterdam University*

CONTRIBUTION TO THE PSYCHOLOGIC AND  
LINGUISTIC VALUE OF MELODY

BY

L. KAISER

The importance of those qualities of speech sounds, which are linguistic and extra linguistic at the same time, is generally admitted. The question to what scientific discipline researches on melody and rhythm belong, can hardly be answered at this moment.

It is known that Van Wyk wished to reserve a special science for this purpose, situated between phonology and phonetics. Duyker described them among the extra lingual phenomena. Kainz called them primitive means, whereas the author chose the term biological component. Trubetzkoy in his last publication considered rhythm and melody to belong to the field of phonology. William Stern used the multiplicity of speech research as an argument for the foundation of a new science to be called personalistics.

As a whole there is but little known concerning the essential features of speech melody. Smits van Waesberghe demonstrated that melody in speech is inseparably bound to melody in music, in the first place in songs.

As during the last year others were engaged in recording speech melody in various cases, looking for the melody patterns that might have a more or less fixed value, I put the question whether a sequence of tones contains an impersonal message.

The impression caused by this message, might have a psychological or a linguistic character, or both at the same time.

The stimuli consisted in 49 sequences of three tones each, all of them beginning with c'. The interval between two tones might be: ascending, equal or descending. Ascending and descending might be: slight (one tone or a half), moderate (a third) and large (a fifth). Miss B. Uylings kindly noted the 49 combinations in a fully arbitrary sequence. Apart from myself six subjects



were listening to the tones, played on an enarmonium or a piano, deprived of any difference in intensity or duration, as much as possible, the mean duration of each tone being about  $\frac{1}{4}$  of a second <sup>1)</sup>.

The judging was done partly independently, partly in mutual deliberation. The female subjects almost invariably judged the value of the three tones by giving sentences of three syllables which had presented themselves to them associatively. One of them often recognized a given melody as a part of a hymn or psalm. The three male subjects judged the value of the melodies from the point of view of their special fields. One of them, a singer, formulated the psychological sphere inherent to each of the small melodies, without any interference of language associations. A linguist especially judged the linguistic possibilities and characteristics contained in the melodies, whereas the third, a psychologist, often succeeded in combining various judgments given by others to a single one, that might be at the bottom of the former.

In all cases a remarkable degree of accord was reached, the divergencies left all regarding the linguistic interpretation of the melodies.

Here may follow a table of the results obtained.

As an illustration to the table (p. 290) a few examples are given of the results in their original form. The expressions have not been translated as this seems hardly possible.

cAc displeasure:			
aanmaakhout,	zie je wel,	waarom niet,	't doet me niets.
cBA seriousness (displeasure):			
ach wat naar,	zoek het maar,	'k weet het niet,	hou toch op.
ccB low spiritedness:			
och hij valt,	dag mevrouw,	het is goed,	complaint.
	(by a servant)	(dull)	
cde encouragement:			
lust je drop,	—	weet je wat,	introduction.
ccF respect:			
het is uit,	't is te laat	't is genoeg,	constatation.

<sup>1)</sup> I thank A. de Graaff, G. L. Meinsma, B. J. de Nooyer, O. E. Steffen, B. Uylings, J. M. de Vries and B. Wilbers for their kind assistance.

# FIRST INTERVAL

	large ascens. (stimulation)	moder. ascens. (cheerfulness)	slight ascens. (quiet cheerfulness)	level (passivity)	slight descens. (sadness)	moder. descens. (sadness resignation)	large descens. (sadness resignation)
large ascens.	P deprivation of contents L question	P personal directedness L question or call	P cheerfulness L question or information	P cheerfulness L question	P peevishness opposition L question	P melancholy L question	P indignation L refusal (!)
moder. asc.	P irritation L half a question	P cheerful surprise L utterance or question	P composedness piousness L question or call	P hesitation L question	P curiosity peevishness L question or answer	P displeasure L unreal question or call	P cheering up promise L utterance
slight asc.	P interest L question or call	P harmony L „weiterweisen“	P encouragement L „weiterweisen“	P surrender L command	P resignation L half a question	P longing L question	P passivity L information
level	P confidence optimism L call	P cheerfulness L question or affirmation	P soothing hesitation L introduction	P admonition L midpart of utterance	P regret L information	P element of willing L command or call	P gloominess L affirmation or answer
slight desc.	P rallying L utterance	P cheerfulness L hardly a question	P reverence L information	P low-spiritedness L utterance, call or question	P sadness L utterance	P resignation L information	P tragic imploring L information
moder. desc.	P kindly teasing L utterance	P encouragement L question or information	P negative sphere L unreal question	P childish sphere L commanding	P resignation L utterance	P depression consolation L information	P resignation admonition L utterance
large desc.	P puzzle teasing L exaggerated intervals	P uncertainty or reaction to it L information	P discontentedness L unreal question or call	P irrevocableness L constataion	P tragic resignation L utterance or question	P resignation fatigue L information	P unresolvedness L utterance

## SECOND INTERVAL

From the table we may draw the following conclusions.

If the first interval shows a large ascension (fifth), the psychologic value is that of stimulation, which gets the character of teasing if a descension follows. The linguistic form is that of a question, if an ascension follows, that of a call if equalness follows, that of an utterance if the following interval descends.

If the first interval consists in a moderate ascension cheerfulness reigns over the whole series, only in the case of a large descension following, getting transferred into uncertainty. The linguistic form is that of a question, if an ascension follows, the slight ascension giving a "weiterweisend" character. If equalness follows, the character is undefined, if descension follows the informative character comes to the foreground.

The first interval being a slight ascension, a quiet cheerfulness, assuming features of piousness, encouragement, etc. is met with. The linguistic interpretation is that of a question, if a large or moderate ascension follows, the slight ascension giving again the "weiterweisend" character. If equalness follows the whole gets the character of an introduction, if descension follows, a call or an information may be recognized.

The melody beginning with two equal tones, seems to bring the element of the will to the foreground, showing itself actively or passively (in the form of a surrender). Linguistically a following large or moderate ascension induces the character of a question or a call. The other cases frequently show the character of a command or a constation, though no full agreement was reached here.

If the first interval shows a large descension, sadness with resignation and fatigue are recognized. Linguistically an utterance or an information is present. Remarkably enough an ascension in the second part was not able to give the whole the character of a question.

The first interval showing a moderate descension, again sadness with resignation appears. In this case a following ascension gives the character of a question, a following descension that of an information, whereas equalness leads to a command or a call.

If the first interval shows a slight descension, again sadness is present, a following ascension marking a certain amount of resistance to it. The following ascension also gives the whole

the value of a question, whereas a following descension leads to the indefinite form indicated as utterance. Only if two equal tones follow an information may be recognized.

It is obvious that the two interpretations are independent more or less, going their own way.

To the psychological interpretation the first interval is meaning much more than the second, which only may add some feature to what has been built up by the first. It seldom occurs that an impression, say cheerfulness, is balanced by a following impression of sadness, though this must be the consequence if the second interval would have a fixed influence independent of its position.

It seems admitted to conclude that the processes leading to an impression of cheerfulness, sadness, etc. are too slow to give way after about half a second, to give the second interval its chance.

The linguistic interpretation showed itself much more dependent from the second interval, the course of which determined on the whole whether a question, an information or only a call or an utterance was to be recognized. Equalness, slight ascension and slight descension coming in the second place gave rise to a set of incomplete and indefinite linguistic forms.

As stated above the unanimity in linguistic interpretation was much less than that in psychological interpretation. Hence it will be necessary to repeat the experiment, directing it more especially to the recognition of the linguistic form. Elimination of the other part of the experiment probably will be neither possible nor necessary.

The two processes being separated in time and space, hardly touch each other. Whereas the psychological resonance to the affective patterns contained in the melody goes forth in the subcortical part of the brain, the recognition of the linguistic forms must be situated in the cortex. As the difference in rapidity of processes taking place in both regions is a well known fact, it needs not wonder that the course of both processes is widely independent, even in so simple a case as that of a melody built up of three tones.

Nevertheless it occurred that both fields showed a certain relation. The melody c d f gave a special impression (of a ship passing at a great distance), which could be combined very well



with the linguistic function, being that of "weiterweisen". It is impossible to determine the value of this single case.

### SUMMARY

In listening to 49 different melodies consisting of two intervals each, it appeared that the first interval determined the psychological (affective) value of the melody, whereas the second interval determined the linguistic form, which might be covered by it.

Seven subjects hardly showed any discordance as to the psychological interpretation, whereas they did so several times in the case of the linguistic interpretation.

# THE UNDERSTANDING OF FACIAL EXPRESSION OF EMOTION

BY

NICO H. FRIJDA

*From the Psychological Laboratory, University of Amsterdam*

## CONTENTS

### PART I: EXPERIMENTAL PROCEDURE

1. Introduction
2. Experimental Material
3. Experimental Setting
4. Problems of Scoring

### PART II: RESULTS AND CONCLUSIONS

5. General quantitative results
6. Some definitions
7. The articulation of the expression-melody
8. The rôle of learning and of knowledge of the situation
9. The apperception of the particular configurations
10. The evolution of expressions.
11. What does facial expression express?
12. About the process of apprehension
  - A. The direct impression
  - B. Formulating the response
  - C. The forms of interpreting
  - D. Reflectory imitation
  - E. Expression and emotion
  - F. The impression-process
  - G. The conditions for the possibility of the process
13. Summary

Appendix and bibliography

## PART I. EXPERIMENTAL PROCEDURE

## 1. INTRODUCTION

Although it has attracted much interest and has been studied over a long period the understanding of emotional expression still is, and forms a problem to much of a puzzle which no satisfactory solution has yet been given. This problem itself is clear and well-known: how is it possible that we know, or at least presume to know, what is going on in the mind of someone else; to know what feelings and emotions move him without his telling us about them, only by watching the subject's face; how, briefly, someone's face can be like an open book to us. The modern authors like Scheler, Klages and the phenomenologists do not get tired in pointing out the reality of this knowledge, giving it even a primary place among any knowledge which can ever be obtained by man. They often describe in romantic terms the possibility of thus penetrating into the souls of our fellow-men. But almost always they have taken this possibility for granted, stating without much discussion that this insight is obtained by a kind of perception, and not by judgment. The experimental conclusions are still controversial. Even in recent publications the reality of this knowledge has been denied (Turham, 1941). About the process giving rise to this understanding, in case it be acknowledged, very little research has been performed.

The present experiment was planned with the modest purpose of testing the significance of evolution of facial expressions for the interpretation. That is to say, to determine the role played by the way they begin, come to their more or less pronounced climax, and decline again. In short, the role of their actual dynamic aspect. To this end, short motion-pictures were shown of two persons whose experiences while being filmed were known to the experimenter. Single frames copied from the films were also presented as still photos, serving as material for comparison. Both series, consisting of 68 films and 75 photos had to be interpreted by forty observers. The experiment was planned and executed at the psychological laboratory of the Amsterdam Municipal University under the direction of Prof. H. C. J. Duyker.

Over and above this limited aim, it was hoped that the experi-

ment might bring forth suggestions for other and deeper attacks on the problems of understanding expression, expression in the restricted sense of emotional facial expression. That is to say, only those facial movements were considered which correspond with a more or less momentary condition of the person, and not with his personality as a more lasting structure. A stricter definition of what is meant by "emotion" has not been attempted. The theory of emotion is of course of primary importance for the problem. But this being still in a rather undefined state, facts and opinions which seemed useful were incorporated, and sometimes tentative distinctions and assumptions were made.

The relative succes in judging photos and films has already been investigated by Dusenbergh and Knowler (1938). I regret that I have only read their paper in abstract. All I know of their results is that they found interpretation of films superior or equal to that of photos depending on the circumstances. Most of the other problems tackled experimentally will be referred to later on. Not, however, the rather frequently repeated question as to the interpretation-efficiency in judging the lower or upper halves or the entire face. The value of this question seems doubtful, and this is also the moral of one of the publications (Dunlap, 1927). Nevertheless, Buzby (1924), Ruckmick (1928), Frois-Wittmann (1930), Hanawalt (1944) and Coleman (1949) still devoted their energies to it. It is not surprising that the invariable result was the superiority of the interpretation of the whole face, and the variable outcome the presumed dominance sometimes of the upper half (Buzby), sometimes of the lower (Dunlap, Ruckmick), sometimes of the first, sometimes of the latter (Hanawalt, Coleman), or denial of any dominance (Frois-Wittmann).

The first interpretation-experiment was performed by Darwin (1872), although his aim was the demonstration of the meaning of the shown expressions. Our problem, i.e. the possibility of correct judgment based on the face alone, was studied by a whole series of authors (Feleky, Buzby, Ruckmick, Landis, F. Allport, Kanner, Langfeld, Gastes, Kline-Johannsen, Hulin & Katz, Munn, Turham). They worked with different kinds of material, and had different methods of obtaining the judgments. Of course, they obtained different quantitative results too, and not much more than that. With the exception of Landis (1929) and



Turham (1941), all of them concluded that correct judgment is possible. Sherman (1928) made the same experiment with films of infants less than 12 days old, and also rejected the idea of being able to interpret the expressions of you and me. Experiments about the process of apprehension of emotional expression proper have not come to the attention of the present writer, unless one wishes to consider the work of Bühler & Hetzer (1928) as such.

## 2. EXPERIMENTAL MATERIAL

The choice of the images to be interpreted depends of course on the purpose of the experiment. To study the process of interpretation as it occurs in daily life, the highest degree of spontaneity of the expression is to be preferred, and this condition was not always fulfilled. The earlier experiments made use of rather schematic drawings (Boring & Titchener, Buzby, Jarden & Fernberger, Langfeld, all with the Piderit-models; Frois-Wittmann), posed photos (Feleky, Kanner, Ruckmick, Frois-Wittmann, Hulin & Katz, Hanawalt), expressions provoked by electrical stimulation (Darwin), posed films (Dusenbergl & Knower), spontaneous photos (Landis, Hanawalt, Turham, Munn) and spontaneous films (Thompson, Coleman). Landis pointed with reason to the fact that work with posed material may lead to distorted results, an expectation borne out by Coleman.

It is easy to obtain candid photos, but it is more difficult to get them unawares, the more so if at the same time, one wishes to control the situation which gives rise to the emotions, and to demand after each exposure a more or less complete introspection on the part of the filmed subject. Those conditions are nevertheless necessary, as will be clear from the scoring method used in this experiment. It is clear also from the unpredictability and intricacy of the subjects' reactions, in a given situation. When in the present experiment offering one of the subjects a box of candy, for instance, this little expense was made to provoke an expression of happy surprise. Instead, she showed a face in which politeness, distrust, and appreciation were mingled in a most complicated way. In order to know the emotional pattern, I needed her report. The effect of superficial introspection or even its absence in many experiments seems to have diminished their value.

Construction of "moving" situation is not difficult, and Binet and Courtier (1898) already have given some fine examples. It is more of a problem to tap the less reactive inner states, and here one depends mostly on the mental wealth and freedom of the subject, and the aptness of the atmosphere. Nevertheless, some films of really spontaneous moments could be made for this experiment. The exposures were taken in the experimenter's private room. Both subjects were acquaintances of his, and of course, they came separately. The subject was seated in an easy chair; so was the experimenter, facing her, at the side of the camera. Moreover there were two big floodlights, while a snorting machine tried to mask the noise of the camera. It was not a very homey situation, and one may ask in how far the reactions and moods of the subjects were free and uninhibited. One need not be too pessimistic about it. Subject A, a female student in psychology, knew the purpose of the experiment. This slightly inhibited her inner and outer reaction, but on the whole, her attitude was sufficiently free. Subject B, a student at an art school, also female, did not understand what it was all about until the very end of the sessions. She too was not altogether very free, and she often showed a good deal of reluctance, not so much because of the camera and lamps, but by reason of deeper personality aspects. But even if both of them had been more artificial in their behaviour, it would not have been too serious. The introspections give account of the reality of their experiences and moreover, their inhibitions are also part of these introspections. Perhaps it only made the series of expressions slightly monotonous.

The subjects' awareness of the fact that films were taken was an inevitable consequence of the lack of better technical equipment. The camera was an ordinary 8 mm-film apparatus which had to be wound up after every two scenes. Release was unnoticed because of a hidden cable-release; the noise of the camera itself sometimes was perceived, sometimes unnoticed because of the above-mentioned masking noise.

The exposures were taken on 8 mm film, during three sessions of about three hours each, at irregular intervals during the conversation, at moments when the experimenter saw something worthwhile to record, or whenever the occasion for one of the "stimuli" like explosion, dirty-smelling substance, or the offer

of the box of candy seemed proper. The images showed only head and shoulders and no objects of importance for guessing the situation. The length of the films varied from 2 to 15 seconds (32 to 240 frames). After each pose, the camera was stopped, and the subject asked for introspection over the foregoing moments or events. This was literally written down. Afterwards the film was cut, every pose separated by a blank strip so as to permit the stopping of the projector during the experiment proper. Sharpness, detail and brightness of the film were satisfactory, of normal amateur quality. From each of the scenes, one or more salient or representative images were selected and enlarged on 5 by 5 cm lantern slides. In projection the screen-images of film and still-shot were of about equal size and brightness. The series consisted of 31 films and 34 photos of subject A, 37 films and 41 photos of subject B, totalling 68 films and 75 photos. In the following, the films will be indicated by the letter C (ciné), the lantern slides by S (still), followed by the numbers figuring on the images.

### 3. EXPERIMENTAL SETTING

Both series, films and photos, were subjected to the judgment of forty observers, 20 male and 20 female, each of them undergoing the experiment separately. They were selected on the basis of their supposed capacity for expressing themselves more clearly in words. On the other hand, the effort was made to keep the group as heterogeneous as was still possible. A positive selection was, however, made, as appears in the numerical results. These show indeed a negative skewness of the distribution of scores-per-observer. Three of the O's had not attended high school; 25 had an academic education or were university students, of which 11 studied psychology, 5 were artists, 4 had administrative functions, and 6 various occupations, as social-worker or housewife. The selection was inevitable, for judging the series was a strenuous affair, almost too much even for some of the persons chosen. The ages varied from 19 to 39, with a mode around 23.

As was said, the O's (observers) came one by one. They were seated at about 1 to 2 meters from the screen, slightly to one side, so as to permit the rays from the projector to fall over the shoulder. The room was dark except for the shielded lamp



of E (experimenter). Of course this darkness was necessary, but it also had a more psychological advantage. It gave the act of looking at the pictures more the character of face-to-face contact and less of a casual and objective and passive perception. It seems useful in similar experiments not to give out the photos into O's hand, but to project them.

When O was seated, the instructions were read: that he had to relate what "went on" in the subject's mind; that he could do it in his own words, and if he liked, by describing the situation; that the poses for the most part were spontaneous, and that not all were taken during ordinary conversation. Eventual questions were answered, the light dimmed, and at once the first series presented. Every O was first offered a photo or film interpreted by E. The succession of the series varied; sometimes the photos first, sometimes the films; sometimes subject A, then subject B. Usually the entire material was shown in two sessions of about one hour and a half to two hours. Each pose was presented only once.

One difficulty must be mentioned in regard to the showing of the still-photos. One has to remember the initial aim, i.e., the comparison of static and dynamic expression. The static image therefore had to be as comparable to its moving counterpart as possible. This pertains to the time of presentation also. The stills were flashes from the film, but to show them as a flash of course is nonsense. To show them at all introduces a special kind of evolution: a static one. This evolution is applied in the case of an expression of laughter, to which it is foreign, as well as to one of quiet listening, where it belongs. On the other hand, it must be presented long enough to be apprehended at all; but also briefly enough to necessitate O to interpret it from his memory-image. For this is what he had to do with the films too. Moreover, time of presentation has an essential influence on the efficiency of judgment. Dusenbergh & Knower (1938) obtained 64 % correct judgments when the time was short (how short, the abstract did not tell), while 81 %, that is, as much as they found on films, when the time of presentation was unlimited. An exposure time of 10" was rather arbitrarily decided upon, being as long as the average film. Some O's found this to be too short, which was intended, but often the answer was given within this time.



*Obtaining the judgments:* The method of obtaining the answers was thus rather simple. The O's had to state in their own words their impressions and opinions as extensively as possible. That means that they gave emotion-names, descriptions of behaviour, of situations, sometimes exclamations the subject might have made. This method is not self-evident. In the literature, it is only sporadically found; in experiments on emotional facial expression surprisingly enough only in Darwin, and besides, more or less, in Kanner. Sometimes the experimenters got their answers in the most simple and undifferentiated way: by having O check a list, one more in number than the number of poses (Coleman). Other ways were: to have them mark one of a list of emotion-names with three times as many names as the number of designs presented (Buzby); a non-verbal method of sorting in groups (Hulin & Katz); matching poses with the same number of emotion-names (Dusenbergh & Knower); giving a list of names which O could choose but did not have to (43 names, Frois-Wittmann; 100, Feleky, Kline-Johanssen); free naming (Landis, 1929, Sherman, Kline-Johanssen, Turham, Munn), sometimes augmented by adjectives and descriptions of the situation (Kanner); also, the suggestion of a name which had to be judged appropriate or not (Jarden & Fernberger). The wide variations not only of material but also of judgment-method makes comparison of the results a doubtful affair.

Here, too, the choice of method depends on the purpose set by the experiment, and in this case, a premium was set upon spontaneity of O. Free description seems to be the best way for approaching the problem of interpretation-capacity and interpretation-proces. The more so where single words more than those descriptions have individually different meanings and do not clearly indicate the intention of the speaker. In this experiment, the answers were written down as literally as possible. This "possible" contains a vague restriction, because it was done by the experimenter, who at the same time had to occupy himself with switches and the stopwatch. However, it boils down to omission of articles, the abbreviation of long stories which denoted E ("the man or woman, the person who just spoke to her..."), irrelevant additions ("Now I think she is being..."), and repetitions.

#### 4. PROBLEMS OF SCORING

After the experiment had been performed, the inert mass of about 5,500 judgments somehow had to be transformed into something easier to handle. Scoring by a scale of "correct" versus "incorrect" is of no use, as has been discussed by Kanner (1931) and Woodworth (1938). "Fear" for a picture expressing „anxious expectation“ certainly is better than, for instance, "enthusiasm". Some more subtle classification is obviously needed.

As has been repeated very often, every inner experience is an original phenomenon, which in certain respects cannot be reduced to components. Emotions, too, are experiences *sui generis*, each being a new and self-sufficient inner event. But nevertheless, relationships exist between the different feelings and sentiments. On this basis, many classifications and distinctions have been made of different kinds and values. One may consider the emotions as ordered along the pain-pleasure scale, or some other Wundtian dimension; one may look for the similarity of the situations which bring them about, like reactive and spontaneous; one can group them according to the similarity of their expressions, or to that of their phenomenal "color".

For this experiment, some grouping of the last kind seemed the most promising, being the most concrete; more so than theoretical distinctions like pain-pleasure which leave many experiences between them. Moreover Kafka (1937) denies any expressive correspondent to this classification, and the present material points to the same conclusion. The basis for scoring here adopted is the degree of similarity between the emotions implied in the complete response of O on the one hand, and in the introspection of the subject on the other.

What is this mysterious similarity? Comparing introspectively "fear" and "anxious expectation" we find "something" common to both of them, in spite of the many differences, and the difference in total condition. In both, we feel a tendency *away* from some object, or whatever it may be. In fear, that object is present, at least in our imagination. If the external circumstances permit, we give in to that tendency "away". In anxious expectation, the object is not yet there; it is felt as possibly coming, and the tendency consists only of a preparation for eventually hurrying from this place. Or another example. Alarm hits one profoundly, in one's vital existence. Astonishment, on the other hand, remains

PLATE I

Subj. A

PLATE I



Subj. B

PLATE II





more in the intellectual realm. But in both emotions there suddenly appears something unknown, an object or person one cannot place in one's world, and for which no appropriate reaction is present.

Both the differences and the similarities mentioned in the examples form part of the experience-complex of the moment. If action were to come about directly, without intermittent voluntary decision or change of attitude, this action would take place in the direction which that tendency indicates. In fear *and* in anxious expectation, I then would run away; in fright and astonishment both, I then would ask what strange thing happens to appear. It makes no difference whether the emotion is conscious, is really felt and realized by the person. If in crossing the street I jump aside for a suddenly approaching motor-car, I have "no time" to be afraid, but I just run, and this running is acting out the tendency "away". The relationship between emotions is thus found in those tendencies and attitudes. That means in relation to the possible consequences of an emotion: the relationships between emotions are similarities of *readiness for action*. And from the point of view of the moved person himself, it means that these relationships are similarities of specific forms of *position towards the object* or thought-content, towards his world, his "behavioural environment".

The simple meaning of this digression: an emotion can be analyzed, notwithstanding its unique character; it can be compared with other emotions, can be more or less similar. Kanner created a scheme, which seems to consist of terms related in this manner. Maybe it was of use to him, having a number of "best terms" at his disposal as judgments. His photos, too, bore the labels of one name, to which he seems to have stuck rather strictly. Such labels are not available for the present material. The only way was to analyze the introspections of each pose in the described manner, and to test every answer on its total or partial similarity with it. Certainly a subjective factor is hidden in this method. The use of words differs from man to man. The reliability of the method could have been investigated by having the answers scored by more than one person. It has not been tried here. After all, the scoring serves only as a means to handle the material and the quantitative results are but the starting point.

The procedure, then, was as follows. All judgments relative to one pose were assembled on one list, for films and photos separately.

A crude analysis of the introspections was made, and a tentative scale set up, in which the required elements for the different scoring-categories were laid down. Kanner used a scoring from 0 to 10. Here the scale ran from 0 to 4, because of the impression that a further refinement would be such in appearance more than in fact. In scoring all 80 answers to one expression, the scale and the analysis of the introspection were constantly changed and corrected. After this, all answers were rescored on the basis of the ultimate scale.

The general criterium for determining the requisites for each scoring-class was the following: 4 or 3, if the principal elements of the emotions were judged correctly; 2 or 1, if there were still some aspects of the analysis in the answer; 0 if nothing but a faint nuance or even less was present. This made the percentages in the different poses more or less comparable, even when a certain expression was on the whole judged with little success. The difference between 3 and 4 is found in the degree of completeness. For 4, all important aspects had to be mentioned; a 3 could be given if only the total atmosphere was correct, at least one of the principal traits being present; a 2, if one could say, "there is something in this formulation", if some important aspect could still be recognized; 1, if only a more or less peripheral aspect was mentioned. No distinction was made between the main directions of attention, the acts of hearing and seeing. It was soon evident that those two were practically not differentiated by the O's. The only exceptions were made for the poses in which the fixed attention was just the attitude of the subject. Sometimes the 0's gave more than one interpretation of the same pose. The scores were in this case averaged.

Example of the scoring: CA—18. (subject instructed to remember a personally disagreeable event. Sentiment remains somewhat external. At the end, a slight shaking of "no".)

- 4: "Just like somebody telling her what happened to somebody she knows, something miserable. At the end, stop it, don't talk about it any more." The displeasure and the externality are both implied, as is the meaning of the shaking.

- 3: "Clearly something disagreeable is said to her and she disagrees." The mood is correct, the externality lacking; shaking not altogether correct.
- 2: "You showed her something she does not like. She finds no answer, somewhat irritated." Feeling-tone not altogether correct, neither is the shaking.
- 1: "Asked if two colours are alike ... no .. no .. no .. yes. Interested."

## PART II. RESULTS AND CONCLUSIONS

### 5. GENERAL QUANTITATIVE RESULTS

From Tables I and II, the principal results are evident.

TABLE I

	SA	CA	SB	CB	S(A+B)	C(A+B)
Number of poses (N)	34	31	41	37	75	68
Number of judgments	1360	1240	1640	1480	3000	2720
Mean score/judgment	1.319	1.994	1.116	1.664	1.224	1.729
Mean score/pose (M)	52.76	77.76	44.64	66.56	48.96	71.86
% S or % C *)	32.4%	48.6%	28.7%	41.1%	30.6%	44.8%
% "4"	6.8	14.0	6.3	10.2	6.6	12.1
% "3"	13.4	20.7	9.4	19.4	11.3	19.9
% "2"	20.0	26.1	17.7	20.5	18.8	23.4
% "1"	26.2	21.6	25.9	26.5	26.0	24.0
% "0"	33.6	17.6	40.7	23.4	37.2	20.6

$$*) \% S \text{ means } \frac{\text{total score photo's}}{\text{max. possible total score}} \times 100$$

$$\% C \text{ means } \frac{\text{total score films}}{\text{max. possible total score}} \times 100$$

TABLE II

	A	B	A+B
Increase in Mean Score			
$(\frac{M_c - M_s}{M_s} \times 100)$	47.4%	49.1%	46.4%
$\frac{\%C - \%S}{\%S} \times 100$	50.0%	43.2%	46.4%
Increase % 4	7.2%	3.9%	5.5%
Increase % 3	7.3%	10.0%	8.6%
Increase % 2	6.1%	2.8%	4.5%
Increase % 1	— 4.6%	0.6%	— 2.0%
Increase % 0	—16.0%	—17.3%	—16.6%

The increase of the film-score over photo-score thus amounts to almost 50 %, which is a clearly significant difference (Subj. A,  $t = 4.0$ , sign. = 0.006 % level; Subj. B,  $t = 3.65$ , sign. = 0.032 % level; after Fisher's Method; v. Milton Smith, 1946). The number of "3" and "4" scores has increased considerably, while "1" and "2" stayed almost the same. The increase thus is extended over the entire range, and not due to a better apprehension of only the crude aspects.

It is not very well possible to estimate in what measure the increase of 50 % is caused by intrinsic factors, e.g., the evolution of the expression. Some photos were slightly darker or less sharp than the films; the memory-aspect worked a bit differently for the photos and had probably less influence; the effect of the exposition-time has been mentioned but its size is unknown. But those factors counteract each other, some favouring the film, others the lantern-slides, so that we may conclude that most of the increase is a consequence of the experimental variable.

Opinions about the value of the score-totals themselves may differ. On the films, the O's obtained 45 % of the possible total score, on the photos, 30 %; the correct responses ("3" and "4") cover 32 % and 17.9 % respectively. (See graphs I and II). Denying the possibility of correct interpretation of facial expression is clearly not justified. The figures are large enough to be highly significant. In what measure they are so is not easy to say, since the probability-curve in the present scoring-scale is no gaussian, but some J-curve. Which one, I do not know, but to such a curve, the distribution of the photoscores do tend indeed, even if not in any way as much as is to be expected from mere chance. It may not be an exaggeration to say that one correct judgment is worth more than ten failures. The chances for "correct" and "incorrect" are far from even; an almost unlimited number of wrong possibilities stands against a very small group of right ones. Besides all this, all technical and experimental factors which influence both film and photo—8 mm-film, fatigue, difficulties of formulation—lower the number of correct judgments. It cannot be emphasized too much that results like these are minimum-results. The more so since authors like Sherman, Landis, and Fernberger do not hesitate to form far-reaching negative conclusions on the basis of their babies, schematic drawings or photos; with methods even cruder than



those of the present experiment—matching photos with names (Sherman), approval of suggested names (Fernberger) and with a very poor or no introspection on the part of the subjects.

The differences in success with which the one or the other of the poses was judged were considerable (See Graphs I and II). The maximum possible scores per pose being 160, they ranged from 4 (2½ %) to 153 (96 %). The standard deviations of the distributions of score/pose are accordingly high; for the photos,  $\sigma = 24.51$  (Subj. A) and 25.30 (Subj. B), for the films, 25.27 and 28.82. The distribution-curves show some positive skewness, meaning that a few poses did extraordinarily well, compared with the series as a whole. Subject B was somewhat more difficult to judge than A, probably by reason of her more reticent behaviour and more complex personality. The Brown-Pearson-correlation of total score per O between subjects A and B was  $r = 0.53 \pm 0.076$ ; total score-per-O between film and photo: for subj. A,  $r = 0.31 \pm 0.096$ , for subj. B,  $r = 0.40 \pm 0.09$ . There thus was some constancy in performance of the different observers.

The distribution of the total scores-per-O shows a distinct negative skewness, the reason for which has been given in § 3. The standard deviation is very narrow, i.e., about 3 sigma. The mean score for men is equal to that of women, both obtaining 215 points. Kanner and Coleman got similar results, while Buzby and Dusenbergh & Knower found women to be superior.

Then there is the experimental factor of practice. Some increase seems indeed to be present, and in a few instances, a remark relative to this was made by an O. Comparing the series presented first with those coming last (films and photos combined), an increase in score of 4 % was shown; not significant ( $t = 0.73$ ). There was a slight influence on judging the photos of preceding judgment of the corresponding films. Photos given first averaged 224 points, given second, 246 points, an increase of 10 %, significant almost to the 0.5 % level. There was probably no influence of photo-judging on that of the films; averages of films preceding and films given after the photos were 300 and 305.5 points respectively, thus giving a non-significant increase of 2 %.

## 6. SOME DEFINITIONS

There are some concepts which might be helpful in making

certain distinctions in the expressive phenomena studied. If the sequence of expression, the totality of their change going on uninterruptedly in time is called the "expression melody", a nice term suggested by a colleague, one might distinguish:

- a: *articulation*. The experiences of a living person form a continuity. Each new experience evolves from or is superimposed upon the preceding. Of course the very concept of continuity implies the artificiality of the words "new experience" and "preceding". But in any case, the stream of consciousness has a certain structure. So has the expression melody. Someone is looking out of a window quite happily—one expression. Then there is the sound of an explosion—a second expression, etc. Here the articulation is clear; one can almost speak of an incision between the two moments. In another case, he is told something unpleasant, and the smile slowly dies on his face—a fluent change, without such a distinct articulation.
- b: *configuration*. This controversial term only denotes, by lack of a better word, the expression proper, the form of the expressive face at a certain moment. The configuration can be static, as it is fixed on a photo: mouth open, eyes widened, brows raised, etc. Or it can be dynamic, only showing itself in time, as in a defensive movement with the head; backward movement of the upper part of the body, turning the head on its vertical axis, etc.
- c: *evolvment*. By this is meant the development of the configurations in time; in connection with static configurations, the term is unambiguous; it is the aspect which only comes to the fore in the film. The eyes *become* widened, remain so for a longer or shorter time, and return to their position of rest suddenly or slowly. The dynamic configurations, too, evolve. A defensive movement can take place quickly or slowly,—fluently or in spurts.
- d: *amplitudo*. Expressions may be similar, differing only in the amount of muscular movement involved. In fright, I may raise my eyebrows slightly or in a very outspoken manner.

An illustration from the musical field may help to clarify these terms. The illustration might even contain more essential similarity. Comparing the expression melody with a piece of music like Ravel's *Boléro*, the articulation is very distinct in the

repetition of the themata, and less distinct within those. The static configuration can be found in the consonance of the instruments, changing from only strings in the beginning to the full orchestra with woodwinds and brasses at the end. The melody itself is a dynamic configuration. The meaning of the evolvment at once jumps into sight, or rather hearing, when one listens to the rendering of the Boléro by Ravel himself, which takes about 20 minutes, and that conducted by Stokowsky in only 15 minutes. The amplitudo, lastly,—it is magnified from repetition to repetition to the fortissimo of the last round.

#### 7. THE ARTICULATION OF THE EXPRESSION-MELODY

Not all films did better than the stills derived from them (See Graphs I and II). In 14 instances, the scores were almost equal, but in 8 cases, the photos scored distinctly higher. This unexpected result, in contrast also with the general trend, asks for explanation, and it is logical to look for that in the dynamic aspects of the film, as compared with the photo. Indeed, a careful analysis shows the reversal to be due to an interference of heterogeneous elements, not appearing in the photo, with the expression proper. Dividing all 68 films into groups on the basis of those dynamic aspects, these eight stubborn poses rear up their heads in two of those groups. The first consist of films in which a non-essential element is visible in the film and not in the photo, like speaking-movements, eye-blinking, chance variations in the direction of the gaze. The second comprises sudden externally caused changes in the expression, like startling after a merry conversation.

Looking at the incorrect responses given to these last films, we find evidence of a confluence of the different parts in *one* answer, or better, in one continuous event or change of mood. Even in other poses of this group where the numerical results are not reversed, the same thing can be found. The above-mentioned example (A-11), an explosion in the middle of a careless conversation, resulting in fright, comprises among its 12 "1" and "2" answers a big majority of conversational reactions: "lively discussion, something is said to which she reacts very actively (?) reaction of disapproval, something of the "ridiculous" in it". Or "You are telling her something, and she does not understand it. Then she suddenly says, 'Oh, yes, that's it!'".



In other poses where there are not so much two discrete parts of expression as two alternating behaviour-tendencies, similar errors are made. B-42 (looking alternately at a frightening object and at the experimenter) elicits the interpretation: "looking at something moving up and down which frightens her". This and the preceding example give the clue to the cause of the errors, and at the same time to an important property inherent to the interpretation-process. In B-42, there is an alternation, be it of the objects or of the subject's direction of interest. In A-11, there is a sudden reaction, be it from the noise or from some remark. The *expression-melody possesses its own articulation* which is interpreted as an articulation of the experience-stream of the subject. A similarity exists between both, as has been emphasized by Klages (1935), Köhler (1929) and Koffka (1928, p. 131; 1935), and under the experimental conditions the observers supposed the same congruence to be present.

These experimental conditions have a distinct effect on the articulation. In the first place, the films have a beginning and an end. Never in daily life do we see such an isolated scene; or rather, we are aware of the fact of the isolation caused by our own late arrival and quick departure. But here it is presented to us as a whole. That means that an artificial articulation has been created, a strongly closed Gestalt which imposes its unity upon the observer. Moreover the situation is not visible on the film. Nor can O hear the explosion or see the experimenter, or the cup with hydrogen-sulfide on A-135. Outside the laboratory, he would have perceived all this as creating an incision in the course of the objective events. It would have structured his apperception of the expressions, too. It would thus have supported his interpretation, support stemming from a knowledge of the situation, but, to be sure, from a very limited aspect of it. It is not the *content* of the situation which is helpful here, but its articulation. In understanding, this is integrated with the articulation of the expression-melody; both are brought into harmony, and together form part of the insight. If this integration is impossible, as happens in daily life, too, a specific reaction ensues. If the expression is structured but not the situation, we at once ask a question, "What ails him?" In the opposite case: "Oh, how strangely quiet he remains." We become surprised, which is the equivalent of non-understanding. As one



of the O's said of CB-12 (thinking, suddenly shaking her head): .... "But why should she all of a sudden say 'No.'?" Here at least the experimental process is not inferior to that of daily life.

In the other group of reversed relation of scores, a similar disturbance is present. Here non-essential or non-expressive elements were the most forceful in determining O's impression, by purely perception-psychological reasons. A kind of figure-ground relationship exists between the light frown on the somewhat stiffly held head of B-60 (effort to understand whispering)—and the wandering gaze. The absence of an object for that gaze was not apparent while its wandering merely added perhaps some color to the concentration. It is for good reason that the poses of this group could have fallen in the group of static expressions, without change in time, had the non-essentials been absent. Expressions which remain for some time are never very pronounced, never have a big amplitudo.

Those non-essentials, too, can only be recognized as such with the help of articulation of the situation, or by its content. Only then do they reveal themselves as non-expressive features, appearing as purposeful acts (speaking), reflexes (blinking) or more or less fortuitous movements (wandering gaze).

Articulation-mistakes are spread throughout all reports, and may even support the interpretation. There were two poses consisting of rather discrete and barely recognizable parts, but which, taken together, pointed in the same direction. The film/photo-ratio here became 3 : 1 as against the average of  $1\frac{1}{2}$  : 1. In other cases, expression-melodies played at the same time may be fused by lack of separation marks, or, on the contrary, may be split up into incorrect parts. As in SB-42-2 (Somewhat frightened, but with a forced smile of indifference): "As if she is looking at something pretty, pleasant. Here, too, the trait of irony and defense." The smile and the defense are attached to individual causes instead of together forming a defensive smile. It must be added that most errors of this kind probably originate during the phase of verbalizing the impression.

As was said before, in all those instances, the situation is an important aid in understanding the expressions, but only by means of its articulation. In others of the illustrations mentioned, the situation itself seemed to have been of importance. But here, too, it was only a rather formal aspect of it: the fact of the

direction of attention on an object or of the turning away from an object, i.e., Kafka's "characteristics of content" (Inhaltliche Merkmale) of the expression. It is at once seen whether the gaze goes to an object, or rests on empty space when the eventual objects are also visible.

To summarize, we found that apprehension of expressive movements is not a matter of perception of isolated expressions as they are described in most experiments and in the descriptive books by Darwin and Krukenberg. It is a living apprehension of expressive events, possessing their own articulation, being a strong or weak Gestalt in the sense of mutual determination or of independence of the structural parts. This apprehension is supported by the integration of this articulation with that of the objective events, of the situation. This is the reason that a film-experiment, built up with "pure" poses, giving only one movement each, is incomplete as an experiment.

#### 8. THE ROLE OF LEARNING AND OF KNOWLEDGE OF THE SITUATION

The influence mentioned in the last paragraph is but one of ways in which the situation may assist in understanding the expressions.

In the reports of the experiment, simple answers like "gay" or "intent" are relatively rare. Considerably more often, the subject was placed in a more or less concretely described situation. B-55 (continuously pulling a cord), "Just like she sees something very nasty. Something of contempt in it, a bit tense." CB-57 (quick pull at the cord), "It may be she gives a hard push on something with her foot. Anyway, a moment of effort or pain." These examples are not the most extensive kind, but translated into purely psychological terminology, the first would run: "anger and contempt, with intention of communicating it"; the second, "brief bodily effort or pain". The details of the "stimulus" (remark) or action (push with the foot) are inventions of the O's. They are, of course, incorrect, but the emotional content is nevertheless sharply indicated.

CA-12. "Here she is busy with those electrical things, as if she is trying something, then it hurts." The O did not know a Rumkorff had indeed been used. The apprehension and also the memory of similar situations must have been unusually sharp to

react to this difficult film in this way. But compare it with another answer: "Just as if she is telling that she is hurt somewhere, when you touch it, that someone is feeling it, that it suddenly hurts, that you start suddenly." Almost the exact same feeling is coupled with a totally different action or event.

The fact that descriptions of this kind are so much more frequent than simple naming of the emotions, is not only a consequence of the instruction to formulate the expressions. It is certainly easier to express them in this way than to search for emotion-names and to give each of these their proper weight. But often the answer came so quickly, that it looks as if the impression is at once transformed into a view of the subject-in-a-situation. The instructions did not ask for concretization, and the introspections of the observers do not give any indication that they consciously tried to do this. Rather the contrary is true. Some O's worked directly from a concrete attitude. "Sometimes I saw at once what it was, on the basis of the first impression. But sometimes I had to try hard. The answer came when I tried to imagine myself in a situation which checked with the face." Another O: "...With the films, I placed myself in the situation, 'what would happen?'" Ruckmick (1921, p. 34), too, mentions that visualizing the entire person and the situation were frequently used as aids while searching for the right term.

How these concretizations came about cannot be explained by reasoning by analogy. Perception of a face, solely by itself, to which a former situation can be applied, does not take place. How the concretizations do arise is not clear from the material, and I can only give a few suggestions.

CA-15. "Can I really have those Nylon-stockings? No, you don't mean it!.... pleasant surprise, some doubt". This O was not the only one who interpreted this film as pleased greed. But in preparing the experiment, E had told him the idea of offering the subject a pair of nylon-stockings as an emotional stimulus. Is it by chance that just here this memory comes to consciousness (eyes widened, faint smile, lively movements of the upper part of the body)? In this case, it is obvious why the memory is present, and it shows that the impression of the subject's experience carries the situation in its train. It shows also that *the memory may stem from a non-optical former experience*, in this case a verbal one.



CA-17 (tension, resistance, fixed and anxious attention): O 1 judges: "As if looking at something with fixed attention, a game or something, tense, two cars which almost get into collision, but nothing happens. Then she says, 'Gosh, who would do anything so stupid!'" Some time later, this O was questioned about the response, and it is to be regretted that this was not done more frequently. "I got the impression that she sat here (room of E) before the window, and that all of a sudden I saw two cars coming over the bridge". (Did you ever see a similar expression?): "I think that everyone does it in that way, don't you? I think that I would do it that way too. I did not see it in anyone else in particular." In this example, there is no concrete memory. O keeps to the generality of "everyone" on the one hand, and on the other, to the possibility of her own similar reaction. It seems as if in this case, the concretizations stem from the *experience-possibilities* of the observer. That this may be of importance also appears from a remark of another O, saying that she felt her judgments becoming less varied because the succession of the films made her think that "that girl cannot change her moods so quickly".

In daily life, the concretizations, of course, are given. They bring about not the interpretation proper, but the concrete content of it, which is one aspect of the emotion but not all of it.

In the two examples of judgments on CA-12 (the induction-apparatus), "something" was constant, but the supposed situation differed. And not only the situation. The emotion consists of more than that "something". In its fullness, it comprises elements of the particular circumstances at hand. An electric shock is disagreeable, but this weak current did not really hurt. It is the same, for instance, with CA-17. The tension while waiting for the machine to start again is more personal, it involves the subject more than does the spectacle of two cars almost crashing into each other, seen from a safe distance. The one unpleasant emotion is not the same as the other, even if both give rise to a feeling of "dejection", "retiring", or whatever it may be. The fullness of feeling is determined in part by the concrete situation at hand and insofar indeed, knowledge of the situation is necessary for complete understanding.

This influence of the situation on the apprehension-process can be represented as a set of the observer. Somehow this acts



selectively for him on the possible emotions which the subject might still have. Some are excluded by means of it. Very clearly this set is seen in the confusion of the different levels of behaviour, the sensual and the more spiritual. Notwithstanding the remark in the instructions that not all the films were taken in conversational situations, these last were nearer the world of thoughts of the observers than pulling a string or smelling hydrogen-sulfide. There was some set for conversation, which only expressions of high amplitude generally broke down. (As in A-13-2; B-57). It is because of this that the reports of CB-55 (constant pulling of a cord) present not one "4", not one answer in the field of physical effort; but 33 % of the answers contained some *mental* effort, not a bad result, taken by itself. Often the atmosphere of the real situation was keenly perceived, and commentaries were given after telling O what it represented, that he "would never have thought of anything like that". This attitude was sometimes reinforced by lack of articulation, so a pleasant chat at the beginning of CA-135 changed the disgust-reaction into disapproval of a dirty joke or to another kind of mental resistance in 25 % of the answers; or the explosion in CA-11 became a sharp remark.

The knowledge of the situation or its experimental correlate, artificial articulation and prejudiced set clearly *do not determine* the interpretation. They give it colour, and in daily life they limit the field of possible emotions of the subject. Already, at this point of the discussion, we may refute the opinions of Landis and of Turham, which maintain that interpretation rests altogether on knowledge of the situation.

A more direct influence of learning, in the sense of reasoning by analogy, certainly exists. But its occurrence is rare and has little result, which is shown by many erroneous judgments. Some of them are conclusions more from purposive acts than from expressions, as, for instance, in CA-14 (thoughtful speaking): "She is constantly nodding 'yes', so something is said to her with which she agrees." Sometimes the impression of personality has its influence on the interpretation, and this clearly indicates the limited value of analogies: SA-14, O 3: "If I would look this way, it would be defensive. But I don't think it to be the case with her." A more positive effect is probably obtained by the fixed association of a few outspoken expressions with certain emotions.

The extremely high scores of laughter (B-28) and of fright (A-13) can be accounted for in this way. But it may also become negative, as in the coupling of a smile with pleasure. Most of subject B's smiles were defensive and many O's made errors of perception (see next paragraph), because the conventional association hampered their penetration into the more subtle details of the expression.

Acquaintance with the subjects did not appear to help the interpretation. The average score (film + photo) of the O's who knew subject A was 105.6; the other O's scored just as much. For subject B, the figures are 107.2, and 108.8 respectively. This in contrast with the opinion of Landis (1929) and of Thompson (1941, p. 32). Thompson, however, had to do with a rather special kind of former experience, e.g., with blind versus seeing children.

Thus far the importance of what may be called outer experience. But there is another kind of "learning" of importance for interpretation: the psychical (not necessarily psychological) experience. Inner maturity and differentiation, the degree of consciousness of one's own feelings and motivations is of more significance than having met frightened or happy persons before.

I just mentioned the defensive smiles, judged to imply amusement or something of that order. From there, a continuous range of possible interpretations runs via "pleasure but not uninhibited" to the right shade. The incidence of more or less deep insight into those expressions varies from person to person, and the increase in defensive responses seemed to parallel O's inner differentiation, the degree of consciousness of his own structure, too. This material offers only a suggestion about the personality-factor in interpretation. O 34, a more or less uncomplicated, or at least a not very introspective woman gives to nine CB-poses featuring a defensive smile, a right answer three times, two times one of "somewhat ironical", and four times "amusement". O 26, in the process of being psychoanalyzed at the time of the experiment, gives the right shade 8 times (and no excessive erroneous interpretations of this kind). The same trend seems to be present in other poses, where a socially adapted expression covers or weakens an underlying emotive state. In general, it was striking how sharply some persons could distinguish and unravel the two components.

It was said that O 26 did not give excessive defensive inter-

pretations. Projection was, contrary perhaps to many expectations, rather infrequent. I only noticed the tendency in two of the forty O's in some answers, not even in many. A very nervous girl introduced tension into various calm poses; and a kind, yielding woman in a happy condition very often saw tender shades where they did not belong.

#### 9. THE APPERCEPTION OF THE PARTICULAR CONFIGURATIONS

The emphasis on the principal result of the experiment, the increase in scores on the films, as compared with the photos, must not obscure the fact of the interpretability of photos themselves. They still got 30 % of the maximum possible scores, or 18 % "right" judgments ("3" and "4"). This is in agreement with the conclusions of Ruckmick, Kanner and Dusenbergh & Knower with posed photos, and with Munn for six of his candid shots. There is at least *some* correct apprehension of static configurations.

Eighteen percent correct answers is not much, perhaps, but here, if anywhere, the numerical results tell only a small part of the story. As the basis for scoring, correspondence in the emotional experiences of the subject and those meant in the responses was chosen. But there are other approaches possible, and starting from correspondence in the expressions themselves is one of them, i.e., the correspondence between the photographed expression and the possible expression meant by the response. And indeed this approach renders an illuminating outcome.

In writing down the protocols, E was struck by the fact that he "understood" almost every incorrect answer. By 'understood' is meant that there was nothing very amazing in the fact that this particular error was made in regard to that particular expression. E realized this, because whenever this understanding did not appear, he automatically asked, "What do you mean, exactly?" Thus it seemed that even in totally wrong answers, the observer might have reacted selectively to the expression, and gave a judgment which somehow fitted into it. On closer inspection of the protocols, it seemed that indeed almost every answer can be imagined as fitting in this way. In how many this is not the case is not easy to tell; in the 5700 answers I only once was struck by it, i.e., on SA-15 (very tense expectation, impatient, some slight amusement too): "Very personal things are being said to her;



she reacts very timidly and very happily." Here I cannot very easily find correspondence.

Some causes of misinterpretation have already been mentioned, which nevertheless start correctly from the expression. Incorrect articulation, concretization errors, etc. There are other errors which are real, but nevertheless do not necessarily impair the principle of the fixed significance of a certain configuration. They are the errors of perception and the errors of evolution.

"Errors of perception" refer to those responses which give evidence that the observer only reckoned with certain features of the face, leaving out or under- or overestimating certain others. Or, at least, these answers can quite easily be understood as correct answers to a photo which represents an expression differing in the said respects, but preserving the general form of the photo judged. One can imagine them as given due to the fact that a certain tension around the mouth has not been seen or that, on the contrary, it has received too much emphasis, or that a certain smile has been judged to be too pronounced or too little, but that it is nevertheless the judgment of a smile. In reading the reports, most of the errors can be accounted for in this way. The surmise of the perceptive character of these errors is supported by the fact that the features which may be thought of as actually being noticed by the observer always are the most outspoken and conspicuous. Only the finer shades and traits have been neglected. The erroneous answer is not concocted out of the blue, nor is it the result of some action of the real expressive content. It is based on a defective, too global perception. To the incomplete perceptive content, an in itself correct expressive content is attached.

This contention may be illustrated by all the zero-answers on two of the photos. These two have not been chosen by reason of their particularly broad or narrow meaning, or whatever prejudicing quality they might contain.

SA-23 was taken while the subject was asked to look at a photo, with the instructions to "try and get the atmosphere". Introspection reads: "You said something about atmosphere, and I had set myself to it. I got it at once. Yes, I must go on looking. In the beginning, I was really interested, but afterwards, I was just looking, a bit forcedly; you told me to." The photo was taken from a moment, when the forced attitude had already started. The faint smile is the exponent of trying to look interested, the somewhat



stiff and tense position of mouth and head comes from the attitude of "You told me to".

The numbers refer to the observers.

- 1: "Is far away in her thoughts" (tension in head-position not seen; rest might be correct).
- 6: "Relaxed listening, kind mood" (smile overestimated, tension not seen; smile as such and head-position correct).
- 8: "Dreaming, passing thoughts, thinks perhaps about nothing. Floating on her mood. Something melancholic in it." (Tension not perceived; direction of gaze incorrect; smile as such and head-position correct).
- 13: "Thoughtful, a bit depressed, a bit questioning. (?) listening" (smile overlooked, tension correct).
- 17: "Here she has fun over something, but she is wholly involved in it. Not fun over something objective; some tenderness in it too (?) listening, looking at you" (tension not seen, smile correct but overestimated).
- 19: "I do not know".
- 21: "Listening to what is said" (?) — (no specification, therefore zero).
- 23: "Again that thinking, listening (?) some goodwill (as 17).
- 24: "Listening to something which interests her. Serious mood" (smile not seen or underestimated; unaware of tension).
- 27: "Certainly active in some way. Quiet, but active. Kind, serious, objective atmosphere" (as 17).
- 30: "Very interested" (tension in mouth overlooked).
- 35: "Very attentive; kind mood, as in all the others" (as 17).
- 37: "Interested, listening to someone (as 30).
- 38: "Here the mood is very pleasant, she hears something which affects her in a pleasant way" (as 17).

This photo has been selected especially because it was rather unpronounced in character. In the first place, one is struck by the absence of any answer, mentioning the more active and colourful or vehement emotions. This indeed would have demanded a more pronounced expression. In the second place, it is striking that the presence of a very faint smile, and the attitude of the head have always been accounted for. The errors lie mainly in exaggeration of that smile, and above all, in not noticing the tension in the smile and in the position of the head.

SB-17-2: The induction-apparatus was working once more. The subject said: "Damn it". Introspection: "It hurt. I did not like it. Anxious waiting too; I did not know what else would come. I wanted you to stop it: I did not want it to become stronger. Rather unpleasant feeling. I also was afraid of oppression, a feeling of being trapped, a rotten feeling. Also something of 'Well, I have to go through it, it is part of the game'" The photo is taken during a movement of turning away.

- 3: "I don't know".

- 5: "Kind, and talking with you benevolently" (position of mouth incorrect; tension underestimated; wrinkles on the forehead not seen; position of eyes and smile noticed).
- 7: "Openly avowing something very important about herself, which she does not mind the other knowing about. A personal problem she should like to solve with someone else." (tension in mouth, and frown erroneous; eyes and smile correct).
- 9: "Something motherly in it (?) looking with interest at something, a glance full of love" (tension in forehead and in smile not seen; eyes and smile as such, and something of the variation of the smile noticed).
- 8: "Very comfortable, after a walk, with the feet in a bath. Sensuous-delicious feeling" (tension of mouth not seen. Eyes and smile as such correct).
- 12: "Soft expression, a bit proud, too. Touched and proud" (as 9).
- 13: "She takes a sunbath and enjoys the sun" (as 8).
- 16: "One would say that she is looking at something she finds pretty and lovely. Touched" (as 9).
- 21: "Taking notes and attention is drawn by a remark which pleases her" (tension and form of smile not seen; error of evolvment; eyes and smile as such noticed).
- 24: "Very personal conversation; she is feeling rather fine, some coyness. (?) frivolous conversation" (tension of mouth overlooked; eyes and something of the smile-variation seen).
- 25: "As if being photographed, a bit self-conscious. Pleasant mood" (as 24).
- 32: "She likes it, almost happy" (as 9).
- 34: "Pondering about pleasant things" (as 8).
- 35: "Hearing something flattering about herself. Pleased, acting some timidity" (as 24).
- 37: "Reading a nice letter" (as 21).
- 39: "Basking in the sun or something which affects her warmly" (as 13).

Here, too we see the predominance of features, standing out in the configuration, and perhaps blocking efforts to penetrate into the more subtle traits. Especially the position of the eyes, the downward glance, is always accounted for. In many cases, it seems to have been the primary determinant (7, 13, 21, 24, 36, 37). They are errors related to those discussed in § 7, in which non-essential elements influenced the rest. Besides, everywhere the smile has been used, but again the variation from the normal relaxed smile has not, or not sufficiently, been accounted for. In this, as in the preceding image, the principal source of error is a deficient apperception of the tension in the face, of the controversial nature of the expression, and therewith, of the emotion. Not so much the configuration as such, as the degree of tension has been overlooked. Just this tension is visible in only very subtle traits. In the words of one of the O's: "I can imagine

someone else giving a totally different opinion. It often depends on a hardly noticeable trait around the mouth“.

How and why a face, or a feature of it, is seen as tense is not easy to determine. Anyway, it is a property which can be perceived *independently of the expressive value*, in an altogether objective attitude. But most probably this perception is of a special kind, different from that with which we perceive an ordinary Gestalt like a drawing or a circle, and presupposing some coöperation of the observer as a person. It is the more striking that the errors of interpretation are partly based upon this property. The deficient perception lacks not so much “ordinary” attention as sensitivity for the tension-characteristics. Maybe this might prove a cardinal point in the aptitude for understanding expressions.

The general conclusion is that the errors of interpretation are not a consequence of the ambiguity of the expressions, nor of the possibility of correct understanding. This is opposed to Fernberger (1928), Landis (1929) and Turham (1941). On the basis of facial expression alone, one can infer at least a large part of the “subjective” experiences of other humans. The quantitative method of the mentioned authors is inappropriate to reveal the real state of affairs. It is not only the right answers, but above all, the analysis of the errors which proved the possibility of correct understanding.

#### 10. THE EVOLVMENT OF EXPRESSIONS

A general opinion among observers was that the films were easier to judge than the photos. Not only because the data were more abundant, but also because they were more natural. One could sympathize with the films. A photo always is and always will be something “over there”, something you look at and that may affect you, but in which you don't take part. Unintentional imitation, about which many statements were made during the films, was never mentioned in the photo-series. It is essential for the attitude of the observer, and for real communication in daily life, to see the expressions come up and decline, to perceive their change and alternation; in short, not to see expressions, but an expression-melody.

I mentioned the errors of evolvment. Only appearing in the photos, they are the logical consequence of the fact that these

do not show how long the particular expression remains, with what acceleration it spreads over the face, and dies down. In A-11 (fright after careless conversation), the main difference between photo and film scores lies in the number of "1" and "2" on the one hand, "3" and "4" on the other. It was respectively 30 and 3 for the photo, 20 and 12 for the film. Twenty-six of the 30 bad responses on the photo are variations on the theme of "surprise" or "disbelief". And indeed, prolong an expression of not too terrorized fright to ten seconds, and one gets precisely the surprised face. How surprisingly correctly the configuration can be apprehended! Even more so than appears from these figures. In the zero-scored answers to A-11, there is not one "grief", not one "joy", or whatever, except a few of the existing emotions one would like. To be exact, we find ordinary conversation: 2; indignation: 1; question: 2. Or to take another arbitrarily chosen pose, B-136, a somewhat depressed intellectual attention. The zeros: introverted pondering: 7; tired: 3; waiting: 2; listening: 2; melancholy 1. One has only to think of the numerous possibilities not mentioned.

The errors of evolvment, of course, could be found in different degree in almost all the pictures. SB-21 showed introverted fright (eyes closed; head slightly retired; light frown), and offered 22 variations of "profound thought". SA-25-2 gave rise to much "laughter" or "singing", because of the smiling talking-movement, evident as such in the film. The importance of evolvment was shown in a special way by B-51, a little joke which did not constitute part of the quantitative results. It showed a fright reaction almost identical with B-21, but this time filmed with accelerated speed, so it was turned more slowly than natural. Not one intentional gesture of disapproval, almost no fright, but shivering, fatigue, "something disagreeable", without any sharp and sudden sides. And that caused only by slowing down the speed of evolvment by one-third.

The greater naturalness and the presentation of evolvment in the one medium are principal determinants of the difference of almost 50 % in total-scores of film and photo. In one of the evolvment groups alluded to in § 7, that of the poses with rapid course, e.g., of sudden fright, this difference amounts even to 78 %. The average increase would have been even greater had it not been that many configurations bear a mark of the evolv-



ment. An expression with high amplitudo cannot last long, while the subtle emotions never rise briskly, at least as long they are not artificial.

It would be a misconception to think of evolvment as only denoting the variation in speed and duration of occurrence. On the contrary, it covers a wealth of forms. First of all, the amplitudo, being for instance greater in a broad than in a faint smile. Secondly, the quality of fluency. A movement performed in a given time can be executed straight-away; it can be done haltingly or staccato; it can start quickly and end slowly. Moreover, we find the variations around the main direction of development of the expression, which can be nervous, with superfluous fluctuations. We have qualities like starting at once from a former level, or hesitatingly; and at the end of a reactive expression, properties of this kind are very numerous. All these aspects together form the evolvment-Gestalt, perceptible independent from the evolving configuration. In its totality, it gives the observer his impression and gives the expression the character which enables one to distinguish fright from surprise in the film.

One could give a graphical representation of the evolvment. In the example of A-1, it would look as follows:

fright:



surprise:

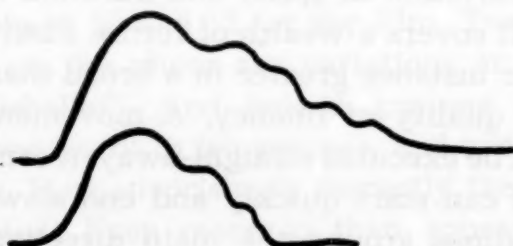


It must be noted that a surprised face, which would be identical with SA-11 is not a spontaneous reaction. Otherwise, the amplitudo could never have reached the given level. Especially the acceleration, the briskness of disappearance, gives the clue to distinguish real expressions from artificial, posed or socially intended gestures. Controlled or voluntary "relaxation" of muscle in general is characterized by the somewhat sudden pull of the antagonists. Everyone who tries to bring his face back to neutrality after an insincere, polite smile will feel how difficult this is.

Very clearly the unfree character of sudden, of too sudden stopping, is illustrated by the laughter of B-28, which was cordial, but died prematurely because of timidity.

uninhibited:

CB-28:



In 22 of the forty judgments, the inhibitions were correctly interpreted, in 36 cases, noticed somehow, and only 4 O's did not mention it. A result attained by few configuration-properties.

But something very important must be added to the graphs above. As was the case with the configurations, apprehension of evolvment is not perception of an objective movement. At least, it is very improbable that it should be so. Most of the terms which denote evolvment have a not entirely objective implication, or are not real visual concepts. E.g., brisk, fluent, nervous, elastic. Evolvment is not only a change of pattern, it is a change of the degree of tension. In what measure the one and the other determine the impression it is difficult to say. But it is hard to imagine that inhibition is apprehended only from the brisk objective movement, and not from the evident conflict of tendencies, the push to express and the brake to keep it hidden. Whatever the case, again as with configurations, tension-evolvment is perceptible as such. But curiously enough, it is not in the same degree visible, independent of expressive content, as was the tension of configurations. On what basis evolvment of tension is perceived is not clear, and certainly it is closely tied up with "material" evolvment. Together, they form a coherent and undifferentiated whole. And thus as a whole, they give life and concreteness to the meaning of configuration too. Small wonder, that the photos gave the O's an often painful sensation of insecurity about the correctness of their responses. Indeed, the average frequency of more than one judgment, given on the same pose, by the same observer, and which had to be scored differently was 0.397 in the films and 0.787 in the photos!

## 11. WHAT DOES FACIAL EXPRESSION EXPRESS?

The foregoing experimental results permit, in the writer's opinion, a conclusion in the important question: Does every facial expression have its own and invariable meaning, and if so, what meaning? Or is each expression representative for a more or less sharply defined group of psychical phenomena? Or even, is it a matter of chance which expression appears at a certain moment? (Landis).

The last contention has already been refuted. It does not account even for the errors made by the observers in their judgment. But the facts reported in the last paragraphs do not leave room for the first contention either. It is not true that to each expression there is one and no other corresponding feeling, sentiment or inner state. Without knowledge of the situation, different levels of functioning are confused, and the concrete colour of feeling is not determined (§ 8). Without data concerning the situational articulation, the expression-articulation remains obscure, and therewith the exact content of the emotion (§ 7). Configurations point to different emotions, like fright and surprise (§ 9), if evolution is not seen. Evolution itself may be common to a number of different emotions (§ 10).

In § 9, we found that in the experiences, whether they formed the real or the supposed basis of correctly perceived configurations, "something" was present in both. This "something" is not the experience in its entirety. A slight electrical shock gave rise to an expression, interpreted as pain, while this interpretation could not be explained as some kind of perceptual error.

A shock and pain are not identical experiences, but both "contain" a tendency of withdrawal and a tenseness. In the disgust of a dirty smelling substance as well as of a morally disapproved remark, one can find rejection and shutting-oneself-off. This connection is generally explained by association. Dumas considers the reactions to sensations on a spiritual level to be derived from those of the sensual; he calls them "transferred mimic" and cites Wundt: "*Le principe de l'association des sensations analogues reçoit une application multiple avec les mouvements mimiques de la bouche et du nez*" (p. 346f). Aside of the necessity of exchanging the term association for one indicating a more intrinsic coherence, the question to be asked

is what exactly must be understood by the analogy of the "sensations", or, put in more modern terms, the experiences.

Discussing the analysis of emotions (§ 5, problems of scoring) we were lead to delineate the "something" of which this analogy consists. It was the similarity or even the identity of "position towards a content". Let us keep clearly separated configuration from evolvment, Strehles figural from his dynamic properties. It then appears, that *the configuration of facial expression is the expression of the persons specific position towards a non-specific content*, that is, of the specifically structured relation he has with an object, person or his entire world. Defined this way, one may indeed say that every facial expression corresponds with one kind of positionality or relation, and with no other. One may state that they do possess a fixed meaning, but one not to be found in the inner experiences as such, in their fullness and concreteness.

The place of this "positionality" in the experience-complex which is called an emotion, may best be made clear by means of the data in the experiment, which lead to defining them. In sensual disgust and in disapproval on a more spiritual level, which, as is well known, have very similar expression, the identical element of rejection is found; of not wanting the object, and wanting to turn away from it, shutting oneself off from it. The similarity is found in an identity of the relation towards the —this time disgusting—content. For after all, a dirty joke, if one does not like them, is "impossible to listen to." One is always directed towards a part of one's world, which is not only a fact, a perceptible object of the geographical environment, but also, and in the first place, the carrier of some value, or valence in Lewin's terminology. In Sartre's essay about the emotions, this is lucidly formulated: "L'émotion est une certaine manière d'apprehender le monde," (p. 30) "elle se pose comme une certaine relation de notre être psychique avec le monde; et cette relation, — ou plutôt la conscience que nous prenons d'elle — n'est pas un lien chaotique entre le moi et L'univers: c'est une structure organisée et descriptible" (p. 16). This apprehension, this relation, is the one appearing in the facial configuration. The direction of an expression can only be described as a relation to the outer world (Buytendijk, p. 373). The constant and objective aspect of the world to which the valence adheres — the electrical stimulus,



the joke or smell—is of no significance for the expression, or at least for understanding its configuration. Hence the addition in the above definition, that the position is one toward a *non-specific* content. The specific content appears only as a totally undefined “presence” which asks for escape from it, for being known better and being placed in the image of one’s world, or which cannot be controlled. Here one encounters the opinions of Klages (p. 126) and of Kafka, that the *content* of an emotion is never visible in the expression. “Nicht *was* getan oder gemacht wird, sondern *wie* etwas getan oder gemacht wird, entscheidet über den Ausdrucksgehalt der Handlung” (Kafka, p. 280) This “*wie*”, this “*how*” the action is performed, is identical with the position which the world forces upon me, or which I adopt myself. Positionality is a fundamental, dynamic aspect of every action and, at least as far as it appears in facial expression, it contains the essential relations. These are indeed very real, even though they are not directly given in introspection.

The standpoint of ordinary introspection is as inappropriate for the theory of expression as for that of emotion. If one would like to consider “fright” an inner experience, one would obtain most paradoxical results. A-11 and B-21 are totally different patterns. The extroverted subject A puts, as it were, a question, she is fixed upon the object which takes her by surprise, and which she want to put into its place. Subjects B finds—not only in the experiment—the highest values in herself. In fright, she shuts herself off from the disagreeable outer world, she does not like to have any intercourse with it, without however, giving up her position by actively fleeing it.

The errors of evolvment in A-11 and B-21 and other poses show that only the configuration is the carrier of this positionality-expression. The mentioned description of B-21 fits also for profound thinking. The subject does not want to be disturbed in doing so. The supposed surprise of A-11, like in fright, brings up the question “what is that there?”. If the experiment had been scored strictly according to the basic relations, the quantitative results would have been still better than they are. SB-27-1 is the pose interpreted worst but one in the entire series. It scored only 14 points on a possible maximum of 160. It was a complicated thing indeed! Miming appreciation (something like “gosh”) while really she only distrusted E’s actions; she did not

know very well where she stood when E offered her a box of candy. Analysed according to the fundamental positions, one would find: a.: appraisal, probing the value, and b.: keeping a safe distance, insecurity. With c.: the simulated positive content of the appraisal. In the 40 answers, we find not less than 32 in which appraisal is somehow mentioned, be it positive or negative. Mostly scepticism, disdain, or sizing up some object or utterance.

The strongest proof that not the experience as such is expressed is given by two introspections of the subjects. A-11, again, was fright. Indeed, the subject herself named it as such, but she began her comment with: "no introspection. First, why does the camera run? Then, emptiness when the explosion sounded". Even in naming one's own emotions, positionality dictates the term. Only under special conditions do the subjective feelings come to consciousness, exactly as in the interpretation of the behaviour of others.

The second example is B-57. The subject was instructed to give a sharp pull to a string attached to the floor. The protocols, besides 25 answers of physical effort and 12 times of mental effort, contain also 16 times "pain" (almost all of them on the film). The expression corresponded completely with seating oneself on a pin or something. But in that case, the expression also is mainly an expression of effort, a momentary pulling oneself together, a moment of biting the lips in suppressing the pain. This certainly does not come to the fore in introspection; it is even almost unattainable for reflection. It is simply done, and the expression has only an indicative connection (Kafka) with pain. It is actually not surprising that Landis (1924) found such divergent expressions in one and the same situations, and even with identical gross introspections.

So much for the configuration. For evolution, too, the general meaning can be determined. What exactly constitutes the difference between amazement and fright? As noted before, it lies in the kind and degree of personal involvement, in the importance the event has for the amazed or frightened person. The high amplitude of expressions of vehement emotions finds therein its regular counterpart. The slow course of subtle emotions, as mentioned before (§ 10) indicates the higher degree of personal distance. In loving, in being happy, in being melancholy, one does not lose oneself in one's feelings, at least not

as long as a little pronounced expression shows that there is not a state of extasy or rapture. One functions on a purely spiritual level. The fluctuating course of a nervous movement sincerely reflects the insecurity about whatever position the subject may stand in. Whereas the wide and uninhibited curves of merry laughter, like those of a big, loosely-suspended pendulum, can indeed only be imagined in quite a free position towards whatever the funny cause may be. Laughter, as a configuration is the expression of a multitude of emotions (Plessner, 1941), but the particular meaning at a given moment can be determined by the dynamic character. The different kinds of tension in the various emotions, elastic, stiffened, expansive, as they are enumerated by Buytendijk, (1948, p. 353) are variants of evolution of tension, and, in one and the same degree of personal involvement, denote the difference in kind thereof. In general it can be said that *degree and kind of personal involvement come to expression in the dynamic evolution*. It may be mentioned that introspective perception of this evolution is closely bound up with the bodily aspects of the emotions, those aspects telling us after Sartre "le sérieux de l'émotion" (p. 41). In all errors with the photos by reason of supposing a more even or longer lasting course (SA-11; SA-135, disgust; SB-45, disgust; SB-57, pull on a string) the change in intensity was accompanied by a lessening of the supposed intensity of the emotions. Disgust became disapproval; introverted fright turned into concentrated thought; pulling changed to effortful concentration. Which position, as meant by the configurations, exists, is indifferent for the meaning of the evolution. And thus, it is of a more general nature than the configuration. The same meaning is implied whether it is an arm or a leg or the face which moves. It gives the same impression whether it is an evolving movement of an inanimate object (at least if with even tension course) or of acoustic phenomena. As was the case with the configurations, evolution shows only an *aspect* of the experience in its fullness, and even both together do not represent the entire experience completely. Hence the eternal discussions which arise about the meaning of musical works which are indeed without any situation, unless equipped with a programmatic title.

It seems that evolution is the more fundamental of the two, as far as understanding is concerned. It enables one to sympa-



thize, to achieve a real contact, even with creatures whose configurations we cannot understand, like animals. The significance of this sympathization was touched on at the beginning of the last paragraph. It is essential for our attitude towards the perceived, and that is why Klages can go so far as to say that every "perception of souls" is perception of movement. Its primary character appears also from the fundamental importance of perception of the phenomenon of self-propagation (Buytendijk, 1948), which seems to be based upon an evolvment-quality.

This experiment has only the emotional expression in view. But one might venture the suggestion that in the expression of personality, a similar relation between configuration and evolvment seems to exist. Is not on the one hand the physiognomy the carrier of "motivations" in the sense of Klages—again in a very abstract sense—so that we cannot say someone is a criminal, but can with more certainty conclude as to his cruelty, ruthlessness, discontent? Is not on the other hand, the dynamic side representative of his formal experience-possibilities, viz. his temperament?

Configuration as well as evolvment, figural as well as dynamic properties, have proved to possess their invariable meanings. Buytendijk's statement (1948, p. 104) that classification of so-called objective characteristics of the isolated movement would assemble very heterogeneous phenomena in one group is not true without specification. The position in relation to the world or oneself, and the kind of involvment are the same with everything, assembled in one group. But actually large differences exist for other functions and the concrete experience.

From the specificity of configurational meaning, there are two seeming exceptions, "purposive action" and "profound experience". The first is an artefact of the experiment and has little to do with apprehension of expressions proper. Sometimes a purposive action is, by reason of deficient articulation (§ 8) considered to be an expression. Real action actually is unintelligible if cut out of its situation, and may by chance resemble expressions, like for instance, closing the eyes on instruction or opening the mouth to speak may resemble thought or amazement.

"Profound experiences" are emotional states of being immersed in one's feeling, of intense listening, and maturing of inner events, or of whatever it may be, in which very much happens inwardly,



and in which the experiences have a rich and almost palpable nature. Poses B-13 (gladness), B-32-1 (listening to a poem; some happiness), B-47 (moment of creativity), B-32-2 (warm nod of understanding) are examples of this. All four of them can boast very bad results. Films and photos both have an average score of 43 points each, i.e. 27 % (average of all poses, 33 %), and nowhere is there a significant difference between photo and film. Only 4.8 % "4" responses were given, while the average for B as a whole was 8.2 %.

Among the errors of B-32-1, there are 23 answers containing variations of "being bored". In B-47, not less than 50 variations of "depressed" are found. The mood being distinctly elated and intent in both. These numbers of one certain error cannot be accidental. They are not a consequence of the profundity of the feelings only in the sense, that the O's might never have thought of these possibilities. These were, on the contrary, thought of at inappropriate moments, and one of the demonstration-poses shown and explained beforehand, moreover, was very similar to B-32-1.

As far as may be concluded from this limited material, it seems that the most profound and lively experiences are not expressed. But this is not surprising after what has been said. The *experiences* are not expressed. It now becomes still more clear why they are not. In speaking of profound inner experiences, one means the conscious experiences, which are reflective phenomena in the terms of Sartre. But a reflected experience is no positionality; it is a certain *content* towards which a positionality exists. It is experienced in an attitude of contemplation. That means at the same time that the person is in a minimum of readiness for action. The little pronounced amplitude of the expressions concerned might already be inferred from this fact. To repeat, the content of the contemplating attitude, in this case the experience proper, never appears in the expression. This shows only the contemplation—or even merely the basic relations present in contemplation. The similarity to boredom or depression is not surprising any more, if considered in this light. There is boredom implying distantiation, aloofness, and there is a grief which is reflection upon the lost object. These elements indeed may be outstanding in contemplation, too.

This attitude of contemplation, of reflection, is rather general,

because it is present in almost all of the more "subtle" or tender emotions, like resignation, understanding, sympathy, happiness or aesthetic admiration. In all of them, the action-component recedes at the expense of the feeling-component proper. Hence the fact that the daily expressions of these moods hardly ever correspond with the descriptions of Darwin and Piderit, with the pictures of Feleky and Ruckmick. These show only the simple, not very profound, but more active forms. Hence, too, that subject B was the more difficult to judge. She was more introverted, with a richer and more colourful emotional life.

It is, moreover, because of this that one is able to predict of many emotions whether they be expressed unambiguously or not. If they are active emotions, then it must be the case. If they have the possibility of becoming an object of reflection, then it is not sure one will notice much while looking at the subject. Timidity always shows, by the simple fact that the word denotes exactly the active form of insecurity. When I am timid, I adopt a very active position, I want to look different from what I am, or to act differently than my first impulse tells me to. I do not want to flee from a situation which is too much for me. Insecurity as a more general concept, can, on the other hand, mean an experience proper, a reflection upon the position—not real and not lived any more—of not being what I want to look like. Then, perhaps, nobody will notice my insecurity. In its extreme form, when I retire from the group of people to think it over, it may not be distinguished from happiness or any other reflective feeling. In the same way, timidity amidst learned company cannot be distinguished from timidity during a first date with a girlfriend. Because, I reiterate, insecurity and happiness, superiority of the savants and the strangeness of a girl are content of the relation. Insecurity was one example. In rage, as the manifest form of aggressiveness, in weeping as the manifest form of grief, one will recognize the same state of affairs.

The spiritual experiences like contrition and shame, jealousy or greed, are not expressed unambiguously. These words, too, signify some general relation in regard to an object which is experienced in a specific way. Unravelling the basic relations, leaves, as determining the expressions, the attitudes like "to place oneself under..."; "being grieved by..." with "to know oneself as hurt"; "to wish to take". The specific rests of experi-

ence are in contrition: "as concerned the value of oneself, or towards someone else"; in jealousy: "by a beloved person", and in greed: "something edible". Indeed, SA-15 (tense and anxious expectation) was often seen as greed, since the subject here also wanted to "take in", as to be prepared for new shocks.

It is exactly these rests of experience that constitute the specific "higher" or spiritual character of the sentiments. And indeed, "Sein geistiges Personzentrum hat jeder Mensch für sich allein" (Scheler, p. 34). These sentiments may be inferred by means of the situation, and only then they can be unambiguously understood. It is not surprising that Ruckmick (p. 33) and Frois-Wittmann (1930) found a much worse result in interpretation of those than in that of the "primary emotions". The results of Munn, who found a better judgment in some and not in other expressions, when the situation was known, and of Turham, who found this knowledge always necessary, may be explained by these facts. The ambiguity of the face alone, instead of implying the reverse, is a support of the thesis of invariable meaning of facial expression, as expression of positionality.

Actually this point of view is not altogether new. The attempt of Plessner (1941) to assemble the different situations which give rise to laughter or weeping under a common denominator, is nothing else than the substitution for a group of heterogeneous experiences, of the relation in regard to two "transition-situations" (Grenzsituationen). Also in the book of Buytendijk there are many remarks, pointing in the same direction. His treatment of the unspecificity of some expressive traits like tension or restlessness (p. 340) and frowning (p. 371) even suggests the existence of levels of specificity. In the subtle book by Lersch the same is indicated for the expression of discrete parts of the face.

As for the connection between emotions—or whatever aspect of them is decisive for the expression—and the form of the expression itself,—Klages attempted to understand the last as a "parable" (Gleichnis) of the purposive action which is possible under the given conditions. Many expressions indeed can be described—and give clues as to their meaning—in terms of action. The frightened face in A-11: a question: what might that be; disgust: I don't want to smell or absorb that; shutting one's senses from the impressions. For the dynamic properties,



for evolvment, the correspondence is clear. The form of the inner movements, the entire time-Gestalt of the emotion is reflected in it. They are "isomorph" in the words of Koffka (1928, p. 130ff) and Köhler. The suddenness of fright, but also of amusement at the point in some joke, is exactly reflected in the sudden facial change. The effort to understand parallels the hesitating way in which the subject nods "yes" in CA-14; in the case of doubt or reluctance the evolvment might have been the same, because there, too, some resistance had to be overcome. As hinted before, the connection between evolvment and emotion lies in the bodily aspect of the last, representing actually the basis for the consciousness of one's own involvment. But all this is not parabolic, it is no similarity to possible action. Klages' attempt at explanation does not altogether hold true even for configuration. In particular, it fails with an expression like the smile. Smiling seems to correspond to an attitude of distance-to-the-world, of "active rest" (Buytendijk, p. 348). The expression comes over one in a state of gladness, which does not want to or is not able to discharge itself in an embrace; in hearing a joke which does not appeal to one altogether, or to the appeal of which one would not dare to surrender completely; in feigning independence and indifference; in superiority which is not important enough to the assumer as to make him haughty. In all these situations one may smile in a more or less pleasant, in a more or less strained, way. The smile is an expression of disposing of action, and as such, it belongs to the group of laughter and weeping. There, too, an impossibility of action seems to exist. But while in laughter and weeping one *cannot* act, and the controls get lost (Plessner, 1941), here one does not *want* to act. For this reason, it is the first indication of the human, voluntary, attitude towards the world. While Klages is not able to account for the fact that just *these* configurations come about, the non-action character of the basic relations may perhaps be brought into harmony with his theory.

The formulations of the responses themselves point in the direction of an apprehension, not of inner experiences, but of positionality. In the beginning of § 8, it has been noticed that simple naming of the emotions, like "loving" or "glad" were rare. In the overwhelming majority of cases the subject was perceived as acting-in-a-situation. And indeed, what



else is positionality, what else is an unreflected emotion, than the readiness to act in a certain way, or even the action itself? So it was in jumping out of the way of the approaching motor-car. The opinion of Scheler, if taken in its strict sense, is not correct: "Dass aber 'Erlebnisse' da sind, das ist uns in den Ausdrucksphänomenen unmittelbar gegeben im Sinne originären 'Wahrnehmens': Wir nehmen die Scham im Erröten wahr, im Lachen die Freude" (p. 6). More correct is the formulation of Katz (1951, p. 82): "Another fallacy of the analogous conclusion hypothesis is the assumption, that an individual pictures himself the subjective experience of others at every possible opportunity.... We displace his expression inward and locate it within him, but not necessarily in his consciousness",—or better still, the statement of K. Bühler (cit. after Koffka, 1928, p. 137): "The child knows absolutely nothing of life and mind, but is acquainted only with purposive events". With adults, it is in general no different. Perception of expression is in the first place perception of *behaviour*, behaviour taken in the wider sense given to it by modern authors like Koffka or Merleau-Ponty. The sharp perception of the inner experience, meant by Scheler and Klages, is the result of a very special attitude, moreover presupposing the coöperation of "higher" processes.

## 12. ABOUT THE PROCESS OF APPREHENSION

The special attitude just mentioned resembles very much the attitude of the observers in this experiment. The instructions asked for explicit understanding and for a formulation of the impressions. In daily life such an attitude is rare, and exists perhaps only in psychodiagnostical praxis. In general, our understanding manifests itself in a more direct way, in which consciousness of feelings, or even of the meaning of perceived behaviour are absent. "Understanding expressive movements means responding to them with adequate non-reflecting movements" (Bühler & Hetzer, 1928). Confronted with a doubting face, we do not ourselves realize the doubt at all, but we simply clarify and repeat our words. In daily life we are set for reactions on an unreflexive perceptual world, on a physiognomic environment.

It goes without saying that the special attitude co-determined the apprehensional process, and that it did so from the start.

Spontaneity, liveliness of contact, and above all, intuition—whatever that may be—and the free and unrestricted rise of impressions will be hindered. But living contact was not altogether absent, judging from the occurring emotional reactions of the O's. Intuition was never completely left out of play, judging from O's introspections. Some facets will be accentuated, some artefacts created, but it seems improbable that the process is qualitatively different in the essential points. Comparison with the few cases in which intuition was really absent makes this fact stand out. Strikingly enough it happened only to two O's who had real contact-problems. With some poses, instead of real interpretations, they gave descriptions of the faces, a reaction which was rare with other O's; for example, CA-32 (listening to a pleasant poem): "At first she smiles, then she laughs more fully, and then she says something." Or CB-42 (tension and anxiety): "She bites her lip!"

There was evidently something wrong with the attitude. There seems to exist a special kind of perception, a specific set towards human beings, or living creatures in general, which gives the basis for the interpretational set proper. It seems that "...a minimum of unspecific sympathy is essential and constitutive for the apprehension of *every* living being—even of the most simple organic movement in contrast to dead movement—as a living being". (Scheler, p. 31).

#### A. THE DIRECT IMPRESSION

Whether he understands expression or not, the observer is somehow affected; he undergoes a certain "Appel" (Kafka).

O 5: CB-56: "Difficult to put into words. It appeals to me, somehow."

O 19: SB-42: "Today I see the expressions very clearly, but I don't know which ones they are. I only perceive the atmosphere. It is curious, that I understand completely the expression as an expression, but do not know where it belongs."

SB-17-1: "I just thought that you could make a perfect psychological portrait-painting without understanding what is going on in the least."

O 40: "I must always try to make my first impression distinct and explicit."

O 26: "When you turn on the light, I receive an impression, a feeling. If I look at it for some more time, this fades away. If I can keep hold of this impression, I follow that lead. Sometimes it escapes, and I must look more sharply."

O 34: "...But sometimes the impression which appealed to me. The thing which matters is: what do you experience, how does it affect you, what does it do to you."

These remarks, for the most part given spontaneously, extracted from the reports, may serve as illustrations. The observer receives an impression at approximately the same time as the image appears on the screen. It is a property of that image situated there and not in the observer himself. The perception-character of expressive meaning is beyond doubt; and whatever explicitations and elaborations must be made before a complete judgment is ready, one has to agree with Scheler and Klages that the direct impression is immediately present and is felt as emanating from the subject's face.

It is difficult to describe this impression more closely. As far as can be determined from O's statements and from the introspection of the present writer, no "experiences" are actually seen. Rather, there is something "behind" or "in" the perceptive content; it seems as if one is confronted with a palpable field of forces, a complex of tensions and strains, of somehow "charged" tendencies upward, downward, advancing or widening. This is not altogether situated in front of the observer; it receives some of its meaning from a resonance within the observer. A component of kinaesthetic coöperation is present while one looks intently at the films and sometimes at the photos. Ruckmick and Coleman have also noticed this. The totality of these undefined forces and "sensations" have a character of evidence, they are very specific, and at the same time an undefinable "going-somewhere", "being-tense" or "being-such". Containing all the elements of understanding, *being* understanding in some manner, it is still a long way from the formulation through a ready and communicable answer. The two sides of the judgment-process seem to have little to do with each other.

#### B. FORMULATING THE RESPONSE

It is easy to see that the long way between impression and formulation may cause incorrect responses to be based upon correct impressions. Maybe all "errors of perception" are the result of blocking the impression by the effort of formulation. Many differences in intensity or ways of combination of various aspects of an emotion may occur as changed. The aspects themselves may become distorted while searching for the right word. Therefore, Kline & Johanssen found a better result when O had to check a list of names, than when he had to choose his own

terms. The process of formulation consists of what Oldfield (p. 44) describes as the successive interaction of concept-schemata with the object-schemata, in this case, the impression, until one of the concept-schemata pulls the right string. The many self-corrections of the O's are nothing but thinking this phase out loud, self-corrections like "she was quietly looking—no—intently looking", or "somewhat disdainful. No, I don't think that is right. A bit funny, and yet a bit tasteless". As was said before, the tentative naming exerts a deforming influence on the impression or on its memory; it fixes the road of search at too early a moment. Often the complexity of the expression forces an active structuration or restructuration on the impression and may bring the O to neglect some aspects. "I gradually hit on a possible meaning. But I don't know whether I did not distort it, if I did not expose one element out of it, exaggerating it, or only paid attention to one detail. I was already glad when I recognized something". (O 4).

#### C. THE FORMS OF THE INTERPRETING

No systematic introspection of the O's was asked for. But the question asked when the series were finished of "how did you do it" offered rather uniform results, and permitted the distinction of three methods in which the response was come by. The degree of difficulty in which any method was used was rather consistent. A fourth method occurred much less frequently and could be inferred from the answers.

1. *The formulated direct impression*: This form has been described under A and B. One cannot distinguish it from simple knowledge, as was supposed in the case of laughter. There are expressions, so well known that such an influence may indeed be assumed.

2. *Reasoning by analogy*: The least important form. Here a real "interpretation" exists as an artefact of the experimental situation. Certain features are linked to certain meanings. These are the cases which might support the older theory were it not that the resulting judgments are mostly of a very low quality. SA-19-1 (intent listening): "She is thinking, judging from the eyes". SB-32-1 (happy listening to a poem): "Not very interested, for she only looks from aside". SB-32-2 (kind smile): "She smiles. She must be amused."



3. *Empathy*, as the equivalent for the German term "sich Einfühlen": The distinction from the direct impression was made by the O's themselves, but none were able to tell much more about it. There is a distinct intention directed towards the subject. O 1(?): "It is more thinking myself into it." (O adopts a sharply directed gaze in saying this). O 6(?): "In general I did not think myself into her, I just looked at her." O 4: "Some films at once had a meaning. With others, I really felt myself into the facial expression, then I gradually came on a possibility." O 29(?): "I let it soak into me. Sometimes I just (!) thought myself into her." O 30: "I tried to immerse myself into her."

The only thing that becomes apparent from these remarks is the special set, of which O is conscious, a distinct intention in the direction of the subject (note the gaze of O 1). This is not clear and does not tell us much about the underlying event.

In the writer's opinion, only supported by his own introspection, this empathy is some "opening of oneself", the constriction of the field of attention to this special object, which creates a strong bond between the observer and object, be it man or stone (v. Weizsäcker). It is an opening, an intentional receptivity for certain "Gestalt-qualities". This attitude does not work towards human beings alone. It can be adopted in contemplating a landscape, a painting, music, or even an-organic movement, which in this way come to life. One may speak of "Gestalt-qualities" because in the mentioned situations, however they may differ, the emotional value, the proper significance, is only completely absorbed, or in the last analysis absorbed at all only if one disposes of every distinction, every active noticing of details. This presumes a certain passivity, the "deep attention" mentioned in the last paragraph (B-32-1). The attitude which concerns us here is not a very simple one, for contemplation as it may properly be called is a very active and intense direction of attention. This explains the sharp gaze of O 1; but the gaze is sharply directed also to avoid eye-movements, which inevitably would lead to analysing. These points of the psychology of perception are, I believe, basic to the understanding of the process of apprehension. When in the remark quoted at the beginning of this paragraph, O 26 noticed that the first impression disappears so rapidly, she noticed the difficulty of retaining the global attitude. Any moment one tends to fall back on ana-

lyzation, distinguishing, in short, on the "intellectual" aspects of perception, probably because these are more useful and most used in everyday activity. A special effort is needed, at least for people who are not artists, to retain the global mode of perception. And indeed, this way one understands still better the independence or perhaps even the contrast between intuition and capacity to formulate the judgments. They spring from totally different sources, and not everyone possesses both to an equal degree, and not everyone can swiftly change sets.

4. *The voluntary imitation:* Mentioned by nearly every O, this seemed to be the last means in arriving at a response, when the other methods had failed. While imitating the expression shown, O waited for an experience to arise. O 3: "I imitate them all (always?), not the full-hearted laughter." O 4: "This one was difficult. I tried to imitate (?), not always, just when they were very difficult." O 27: "Wait a minute, I shall make the same face. Here I have to try hard!...." This method also seems to render a rather poor result, but it may be due to the fact that it was the last way out.

#### D. REFLECTORY IMITATION

SA-30 (intently listening to something personal. Biting her lip): O 40: "I have a tendency to bite my lip too...." SB-137 (making faces), O 21: "Very lively interest, but some pressing away of it. Involuntarily I press my hands around my pipe and stretch them forward."

These two spontaneous statements are given as illustrations of the phenomenon of reflectory imitation. The movement is imitated absolutely unintentionally. By the way, its occurrence in this experiment is a proof that the experimental situation had indeed a real and more or less living character.

This imitation is not linked to the interpretational attitude. On the contrary, it sprang up in a striking way in the experimenter's own case. Sitting behind the film-projection apparatus, he was only intent on switching off the current when each film-sequence had ended. This had to be done quickly; otherwise the next film would come before the lens, and this would spoil its projection. E thereby became sharply fixated upon the image, but not upon it as a human face, only as something which would disappear at a given moment. Nevertheless, he automatically imitated the

facial movements. Not only during the first series did he find himself doing it, but throughout the entire experiment. Although it was rather tiring, the tendency was not easy to resist. Only by consciously relaxing, nonchalantly reclining in his chair, as it were, did he succeed in avoiding his imitations. Then however, a new and surprising reaction came up: he reacted in a very personal way. When, for instance, A-13-2 (fright) appeared, he felt almost guilty, as if he felt he did not sympathize enough with the poor creature deserving it. The phenomenon repeated itself often enough to exclude its explanation as irrelevant chance-occurrence. The connection with the relaxed attitude moreover was evident in introspection.

Reflectory imitation is not linked to expressions proper. The imitating movement in looking at sporting performances (v. the photos in Allport, 1937 and Boring, 1939) or participating when small children eat are well-known, and they too come about involuntarily, and are almost compulsive. Here there is an imitation of vital movements. It is even possible that imitatory impulses spring from seeing inanimate movements; the writer at least more than once perceived kinaesthetic resonance in seeing the nervous movement of sunlight over the water, or the to and fro motion of the branches of a tree. The only special attitude present was one of strong fixation.

The condition for involuntary imitation thus seems to be a very intense direction of attention on the moving object. Guernsey (1928) in her study of imitation, considers "reflectory imitation" of little babies less than one year old an accompanying phenomenon of the more primitive attention, not-conscious, but involuntarily drawn. In that she probably met with the same thing, thereby showing this kind of reaction to expressive and other vital movements to be possible at a tender age. It must be a very primitive process which moreover disappears as soon as a more or less real understanding presents itself in the maturing child (Bühler & Hetzer, 1928).

It would be interesting to study the question as to *what* exactly that imitation consists of. One could easily plan an experiment in which the imitating O's were photographed, and very likely one would not find a true reflection of the imitated face. It goes without saying that in the cases connected with sports, nobody exerts himself as much as the athlete. But during the experi-



ment, too, E had the impression that his facial movements were rather schematic. More striking is the quoted statement of O 21 who mirrored the tension, expressed by the subject with mouth and forehead, by gripping his hands around his pipe. Once E caught himself imitating the pull of indifference around the mouth of A-26, by shrugging his shoulders!

To state plainly the impression one gets from these few observations: not the movement as such is imitated, but the expression itself, the expressed meaning. It would indeed be worth while to investigate this involuntary imitation more profoundly and under a variation of conditions. If the quoted observations are not artefacts, one is forced to conclude, that the impression, evoked by perception of expressive movements, is essential in the apprehension-process, rather than the movements as such. Whatever the case may be, the theory of reasoning by analogy is here disproved once more.

The reflectory character of this imitation must not, however, be considered as being always altogether compulsive. There seems to be a gradual transition, also for O's introspection, from reflectory to voluntary imitation, a transition accompanied by a lessening of the compulsive character and an increase in understanding which may be obtained. At the same time that the compulsiveness changes into voluntary action, the impression or reaction is felt less as springing from the observer and more as emanating from the object seen. In this sequence, E's observation behind the camera, would be the extreme of compulsiveness, and, indeed, he did not notice having experienced any "meaning" in the pictures. While, in accordance with the median position of understanding as it occurs in the experiment, no O mentioned such a degree of compulsiveness of imitation. There exists, moreover, another very typical phenomenon, perfectly fitting into this median position of being part imitation, part on the way to understanding. O rather often came forth with a possible exclamation of the subject, a reaction also mentioned by Ruckmick. An exclamation is an emotional expression, but on the other hand it is something more shaped, carrying more data for understanding than the facial movements. O 28 on SB-18-3, (irritation): "'Now, Nico,now', she says, but what does she mean by that?"; O 37, on SA-26: "... 'How can you think



something like that of me!'. . . . It is difficult to put into words." O 16 on CA-15; "As if she is saying, 'take care'!"

A vocal expression is added to the facial one, and the remaining non-understanding throws some light on involuntary imitation. In the same way as the O with his pipe, as E with his shoulders, O now reacts immediately in the same sense but with a different movement. Imitation thus reveals itself as a symptom of that particular unity of the whole expressive "region". One is reminded of the statement by Duyker (1948, p. 33) that it is impossible to produce a vocal expression, while at the time executing an otherwise directed motor expression. The whole expressive region, as it were, converges into the impression of the observer himself, and it looks like a meandering stream running from the one person into the other, and choosing one channel from a multitude of possibilities: the conscious impression, the vocal, or motor, imitation; and still others, as we shall see. How similar are the exclamations of O's 28 and 37 to the one already cited from O 19; "Today I very clearly see the expressions, but I don't know which one they are. I only perceive the atmosphere." It is not by coincidence that here, too, some of the worst O's gave exclamations of the subject's which as for tone and content seemed very keenly adapted to the real meanings of the photos. What is the use of imitation towards understanding, is an important but a rather obscure question. Klages, while seeing its significance, nevertheless after many a page does not achieve even a clear statement. But perhaps it may be possible to formulate something more about it.

#### E. EXPRESSION AND EMOTION

Thus, there seems to be a contrast in the reactions to a perceived expression: on the one hand, a reflecting reaction of the observer; on the other, the rise of a conscious impression. Between them a gradual transition exists.

This contrast may also be recognized in the relation between the emotions and their own expressions. In the treatment of the "profound" experiences, it was said that if an experience is rich and deeply embedded in personal consciousness, no clear expression can make this experience visible. Similar facts can be found on other levels. Kretschmer's sensitive patients suffered from a poor capacity for releasing their easily aroused feelings

in an overt action or expression. In the Rorschach-test, its author pointed to the inverse relation between motility and the number of kinaesthetic responses of the testee (p. 74). It aids in remembering one's dreams, so full of emotional-motor impulses, if in awakening one remains lying motionless for some minutes (Freud). Overt behaviour, be it purposive or expressive, goes side by side with experiences, which are little accessible to reflection, and the inverse also holds true. This can be formulated in a more general way by saying that a conscious experience exists only when and because no action is performed, no motor release is intended or possible. When and because there is a certain distance to the available aims of action. For the examples of § 12 this is self-evident. There we had the contemplative attitudes of listening to classical music, of a maturing of inner events, of reflection on personal experiences. In the opposite case of primitive fixed attention, there is no place for any distance; there we have exactly the lack of separation of self and environment, the wane of the ego, absorbed by outer events, characteristic of early childhood. In the adequate reaction to expressive behaviour as occurring in ordinary situations, life asks for involvement, but demands also recognition of the other as a human being in his own right, demands "you"-consciousness. Hetero-sympathetic understanding (Kafka) finally again presupposes distance and a reflective, non-involved attitude. One must inevitably think again of Klages' "parable of action" on the one hand, expression on the same level as action, capable of releasing it, and of Sartre's theory of emotion as an action-substitute on the other.

This view may seem incompatible with the often repeated fact that one becomes really angry only if allowing a release to the angry impulse. But this contradiction is only seeming. An emotion may adopt the form of action, or it may result in a conscious experience. Both are manifestations of the „same“ emotion. Now it is possible to repress an emotion, be it in its action-form or in its experience-form. But it is also possible by means of self-control, to transform the active form into the experience-form, if the person withdraws from the situation, and goes to sit in a corner with a pouting lip. Choosing the transformation-way, only the active component is extinguished, the destructive tendency disappears. The experience gains in colour and depth; tension runs high, innervating in a disorganized way

all limbs, and eventually becoming so vehement and unacceptable that we dispose of it with our last resource, losing consciousness. Repressing the emotion as a whole—and this is what the mentioned remark is meant to convey—leaves a much flatter feeling, which, eventually fading altogether, may remain as a dynamic system for a long period. In both cases, it is true, the rage as such had disappeared, but in transformation only because rage can never be a feeling accessible to contemplation. Rage without the destructive tendency is not rage any longer, but hatred. Dumas' passive and active emotions are no independent reactions, but two ways of reacting on the same basic pattern.

#### F. THE IMPRESSION-PROCESS

We found three real methods for arriving at a response: the direct formulation of an impression; empathy; and voluntary imitation. Being directed onto the expressive face, an impression comes about, be it spontaneously or intentionally reinforced by empathy. Reflection lends it shape and judgment-quality. Both together constitute hetero-sympathetic understanding. By reason of the reflective component, it is ontogenetically a rather late acquisition (Kafka, Ch. Bühler, 1928).

Under other conditions, the face gives rise to reflectory imitation or to the unreflexive impression; Kafka's imitative and idio-sympathetic understanding, which are found in children, and also in animals. Which effect is obtained depends not on the expression perceived but on the attitude of the observer. Thanks to his intention, an impression is isolated and reflected upon, or a direct purposive reaction takes place, or either automatically or voluntarily a motor reaction is performed. Furthermore, a fourth effect may result, i.e., transference of emotion, infection of feeling (Scheler), not found in this experiment.

It depends on the intention of the observer which effect is obtained, but the given effect is more or less inevitable and necessary with the given intention. Emotional expression gives rise to a *"process" which while neutral in the first analysis, may realize itself in different ways.*

It is well-known that the reaction to a "stimulus" depends altogether on the condition of the subject (Goldstein, 1934, v. Weizsäcker, 1950). A dog is a different creature to his fellows,



depending on the state of his sexual and other systems. One moment he is comrade to play with, then again he is an enemy in the struggle for the female. The forms of response towards an expression mentioned above, might be distinct ways of reacting, present under different conditions of the observer. But the peculiar fact exists that the different ways of reacting are not independent of each other. The functions of comrade in play, and of enemy, are mutually exclusive. Global and analytic perception, abstract and concrete behaviour, may be superimposed, or the one integrated into the other. But they present themselves as discrete functions. Between reflectory imitation and conscious impression, a mutual exclusiveness also exists, (v. also Scheler, p. 33). However, this exclusiveness is *phenomenally given* during the transition from the one into the other. The ever-present possibility of imitations, is to be found, moreover, in the conscious impression in the form of a kinaesthetic resonance, an imitative tendency. We know and feel that in giving in to that tendency, the experience as such will vanish. It is not true that in imitating, "one lacks the time" or occasion for reflection. The impression, located in the object, simply does not exist any more; upon reflection, we only find our moving body. It is because of these automatic-seeming relations that I am inclined to speak of a "process".

The imitative possibility is present in the impression as a kinaesthetic resonance, and therefore the underlying attitude has a position midway between imitation and cool and unconcerned registration of the impression. The participation, mentioned in the analysis of the direct impression, ensues from this resonance. The "process" thus seems to be the direct result of the expressive content as such, serving as a starting point for experienced affectivity as well as for motor reactions. This is what Klages meant when he stated that the perception of an inner state is preceded by its transference (p. 107). The earlier used metaphor of convergence of expressive stream into the impression is incorrect. It is convergence into this *process*. This is also the description of the mentioned cohesion in the expressive region as shown in the experiments by Duyker, and in the imitation by means of another part of the body than the perceived; and also of the cohesion between motor-system and affectivity, as demonstrated in the work of Flach.



The preceding is in the writer's opinion more than a mere description of well-known facts, because it draws attention to the extremes of imitation, free from any feeling, and of experience, free from kinaesthetic resonance. This also because the daily form of apprehension rests in a midway position in the range of manifestation-possibilities. A range, in which the "amount" of feeling grows as the amount of striving, the "Antriebsgehalt", diminishes (Klages, p. 105).

1. *Reflectory imitation*: this is, as was just said, imitation of expressive content. The necessary attitude is one of primitive, fixed attention, without any distance from the perceived object, which is not really given as a human being or as another animal.

2. *Transference of emotion*: An emotion arises within the observer of a tendency similar to the emotion of his partner. The attitude is very similar to the one found in imitation. It occurs in the same close and unseparated relation. But contrary to the fixed attention there, here man and animal are taken by surprise by the perceived expression, at a moment that a readiness for action is present. The panic-stricken horses mentioned by Klages (p. 351) were walking quietly when one of them started running. In human mob-behaviour, like collective anger, shape is given to an attitude existing towards a common object.

3. *Adequate reaction to a physiognomic character*: I.e. the common daily manifestation form. No experience of the other is consciously represented; the impression is not realized; one only sees an angry face or rather, a menacing face, or one asking for such-and-such a reaction. This reaction follows. The other is more or less given as such, be it not necessarily in an explicit way. The lack of differentiation of the former levels does not exist any more. One is involved in the other or in the less structured total situation. There is a readiness for action, directed towards the other or towards the rest of the situation, towards the cause of the other man's fright, the joke he is laughing about, and which the subject himself will now appreciate. Instead of starting to act, one can initiate an action-substitute, that is to say, becoming emotionated oneself. One feels pity, or becomes afraid that the other will strike out.

4. *The reflectable impression and understanding*: Here, at last, we find the experimental situation. The noticing of the physiognomic character may be the terminal point, or it may

be isolated and related to a "psychical existence". One may direct oneself outwardly or inwardly; in both cases, the distance is conspicuous. A reactive attitude has been replaced by a contemplative one. This can take place with or without any personal involvement, it may amount to observing nothing but a meaningful object in its strictest sense. But the direction of consciousness is the same.

There are many similarities between the reaction to emotional expression and emotion itself. One may receive an impression or perform a motor act; one may feel frightened or simply jump away. Emotion, too, is a kind of "process", or a state of the person, which can manifest itself in a number of ways. For both, the formulation of Koffka holds good (1935, p. 405): i.e., "Emotional behavior will, in our theory, be considered as the dynamics of those . . . forces; conscious emotion as the manifest aspect of these dynamics."

#### G. THE CONDITIONS FOR THE POSSIBILITY OF THE PROCESS

The principal difference between the emotionally motivated perception in general and the perception of expression resides in the genesis of the meaning of the objects, of their "valence". One avoids a dropping stone or a venomous snake because one once learned that they might strike or hurt. The child once burnt shuns the fire. But a happy face does not appear happy because it was seen before, at the same time, for instance, with a generous gift. An explanation by experience is possible for the first case, not, or only in part, for the second. The phenomenology of the direct impression shows the joy—or whatever basic relation may be at the root of this feeling—to be experienced in the object, to be a quality of the person met. However, a decisive criticism of an explanation by experience is herewith not given. Even the fire has gained a "face", a physiognomic quality for the child. Almost all words are for the contemporary human quite arbitrary signs, but he too sees in them what they signify. They contain for the immediate experience, the signified object (Piaget). Thus this objection of Scheler against the theory of reasoning by analogy does not hold. But in this experiment, use of memories appeared only to disturb and to schematize, except perhaps for the photos of laughter. Similarly, the special role of tension-

differences asks for another approach. It is difficult to imagine how associations would cling to these subtle traits, which moreover would seem to entail some kind of participation from the side of the observer. And, lastly, a consideration of the biological importance of expression-understanding, and the minimal probability of so differentiated a memory, especially in animals, and children, lead to the same conclusion. Still another point is the case of understanding expressions one never saw before.

It is necessary to accept the idea of an original apprehension of expression, at least partly independent from former experiences. But one need also come to terms with Scheler and Klages, who with reason never tire of assuring us that in apprehending an expression, one does not see the expression as such. One does not see, that is to say one does not realize oneself, that the position of the eyes, the curve of the mouth are such, the forehead wrinkled, the head fiercely lifted. It presents a problem, that structures which are not noticed nevertheless exert a certain influence. Structures which even *cannot* be noticed, for in an analytical attitude, there is no direct impression.

Fortunately, similar phenomena exist in different fields, and have been better investigated there. Anyone without musical culture, not knowing anything about thirds or leading tones, can distinguish a melody in major or minor. For the impression created by a melody, the term "Gestaltquality" has been in use since Ehrenfels. This implies a general characteristic of the total structure, which can be apprehended in one "perceptive act", in an intention on a Form-whole, not on a "structural-whole" (Révész). We have met this Gestaltquality before, as we have the importance of this intention, in discussing what was called empathy. There it seemed as if this attitude consisted of a voluntary detachment from the analytical kind of perception, in order to let the Gestalt as a whole penetrate us. Every real Gestalt is apprehended in one "act" without the constituting elements as such rising to consciousness, whereas their change or disappearance can thoroughly transform the impression.

It goes without saying that the face is a Gestalt. But recognition of a smile as a smile, of an angry face in a thin or a fullmoon exterior, implies the possibility of transposition. Thus it implies, at least in principle, that one can describe the conditions of the Gestalt-quality, just as they can be described in the melody by



the total of all intervals, and the place of major and minor thirds, the presence of a leading tone; or in a circle, in the definition of the constant distance to a certain point, the center.

This kind of description seems to be possible for the expressive face. Attention often has been drawn to the general characteristics of expressive movements, be they of the face or of the entire body. Grief tends to appear in a downward, hanging, drooping quality; joy in expansiveness, a tendency upward and outward; surprise through stretching in the vertical dimension, an opening-up. For evolution, analogous features were mentioned in the section concerned. Although the dynamics of an elated movement have to be described in many words, they can be apprehended in one act.

Each Gestalt-quality which may be thus isolated might correspond to a certain emotion, or better, to a basic position and involvement-form respectively. It may be so with the expressive person, as indeed it is, judging by the work of Flach. It may be so with the observer of expressions, too.

This is indeed the theory advanced by Köhler (1929), and Koffka (1928; 1935). They point out the fact, of which there is no doubt, that perceptual impressions originally consist not only of material properties, considered until then as the only real constituents of perception, but also of affective properties. These latter are as real, are as objective or as subjective as the former, and spring from the same automatic and constructive process. This theory may even be elaborated, and made to contain in the Gestaltprocess not only the germ of certain affective experiences, but of motor innervation, thus harboring the entire apprehensional process. Just as sometimes from the Gestaltprocess a real global perception may result and sometimes perception of "sensations"—enabling the psychologists of the last century to construct their theories—just as, dependent on set, the motor, the kinaesthetic, or the experiential side may dominate.

This theory would require a close and fixed relation between emotion and expressive movement, and actually this is just what we found in this experiment, as far as apprehension of expression was concerned, and was demonstrated by Flach and by Duyker for the production. If it is right, it would have as a consequence the fact that every movement, or at least every movement as abstracted from its specific carrier, be it mouth or leg, would



evoke the same affective impression. This appears to be partly the case. In an adequate attitude, one does perceive physiognomic properties in the behaviour of a beetle, a tree or a steamship. "Innerhalb dieser universalen Welt des Schauens der Seele aller Dinge ist das Ausdrucksverstehen gelegen. . . . Unsere Empfindlichkeit für Formen und Bewegungen ist Voraussetzung um Ausdruck überhaupt wahrzunehmen" (Jaspers, p. 216). This attitude to penetrate into "Seelen", into core and spirit of all things, has been found already in the set for global perception, for Gestalt-qualities, as present, for instance, in the attitude of empathy.

However, things are not so simple as they seem here. Movements of steamships and beetles correspond only to some of the possible facial expressive movements, and it has not been implied that every expression finds its counterpart in the animal- or non-animated world. There is an aspect in expression we found to be of great importance in apprehension, of a decisive influence indeed in the "errors of perception" and in the meaning of evolvment. I mean the states of tension and their configuration and evolvment. Direct apperception of tension plays a paramount role in apprehension, and said more generally, the Gestalt-psychological theory suffers from being passivistic. In seeing expression, one is involved; in a certain way one is participating in the movement. And up till now, it has not seemed possible to describe the face, to describe the features which represent the facial strain apart from their being tense. It is a fact that we can immediately state whether a face is active or at rest, without in the least being directed onto the expressive qualities. The general descriptions given of the expressions of emotion, the isomorphism of both, all too often contain affective components.

We must choose between possibilities. Either the Gestalt-quality of configuration and evolvment gives rise to a "Gestalt-process", containing the germ for kinaesthetic impression and participation, flowing only from the figural properties of the configuration and evolvment. Or the Gestalt-quality has already been determined by participation with the specific tension-elements. Tension-elements, which are specifically human or animal, specific for the mouth and for the forehead, and exerting their influence as such. Landis actually pointed out the correspondence between the degree of tension and the judgments of

the degree of expressivity of a face. Once again, it is not certain, but it is doubtful that the apprehension of tension be only determined by the geometrical pattern of the face; rather it is suggested in § 10, by a specific sensibility, springing from motor participation in the observer.

Another question is whether this sensibility is indeed specific, or whether it is a special case of motor-participation in every perception. Pálágyi spoke of virtual movements, performed with every perceptive act. Lhermitte states that with disturbances in the body-scheme, with periferal vestibular affections, as well as with the central "total asomatognosia" the entire optical field is disorganized (1939, p. 148ff). Von Weizsäcker and his collaborators extensively studied the "intertwining" (*Verschränkung*) of movement and perception, and in this connection, he makes the statement that in the study of biological performances, we meet with the mutual exclusion of perception and motion. (1950, p. 21), which reminds us of the state of affairs in apprehension of expression.

Some quite different facts need mention. Children, even very small ones pay special attention to the eyes of adults (Bühler & Hetzer, 1928; Gesell, 1928). Köhler noticed an original affective reaction towards eyes in chimpanzees (1919). These phenomena must, of course, be explained in a totally different way, if they do not lend themselves to reduction to experience or anything else. For the moment the riddle of the apprehensional process must be left unsolved and un-understood. In particular, the problem of apprehension of tension and its significance for understanding expression. This is a problem which must be extensively investigated.

#### SUMMARY

The object of the foregoing experiment is the study of some of the problems connected with the apprehension of facial expression of emotions, and primarily the influence of the dynamic aspect of these expressions.

Two persons have been filmed while being placed in emotional situations, most of them arranged by the experimenter. 65 film-sequences of some seconds each were obtained, from which the culminating point of the expressions was reproduced on still

photos. Photos and films were shown to 40 observers who received the instruction to interpret them. The interpretations were scored on a 5-point scale according to their correspondence with the real emotion, as known from the introspections of the filmed subjects, completed by the experimenters knowledge of the situation. The results permit, first of all, to conclude as to the possibility of correct understanding of facial expression. Second, that the results on the film were distinctly superior to those on the photos. The sum of the points accorded, expressed as a percentage of the possible maxima, amounted to 44.0 % on the films, 30.6 % on the photos. The difference is significant to the 0.04 level.

On the basis of the results, we were able to state the existence of an unambiguous and invariable meaning, appertaining to the different aspects of facial expression. Configuration, the figural and static aspect, revealed itself as expressing the specifically structured relation with the outer or inner world, the "positionality" of the person during the emotion, and not of the experienced emotion in its entirety. Evolvment, the dynamic side, seemed to represent the kind of involvement, the personal interest the emotionated person has in the situation, the way he is really touched or free of it.

By means of analysis of the errors made in judgment, it appeared that interpretation is indeed possible. It seemed that the idea of being able to get some insight into someone's inner life, only by looking at his face, is no unreal one. The errors made are no indications of the contrary, but manifestations of deficient perception of a special kind. All erroneous interpretations can be explained if considered as interpretations, given to photos with the same gross characteristics as the photo really shown, but lacking some of the subtler traits. Moreover, we were able to define more precisely the rôle of knowledge of the situation in interpretation. It did not constitute the interpretation, but supported it, by showing the concrete content of the positionality, and by articulating the expressive events.

Understanding, as it occurs in daily life or in the experiment, seemed to be a particular manifestation of the effect, exerted by expressions, an effect which may give rise to other phenomena like reflectory imitation, transference of feeling or an adequate reaction of the observer. Which reaction is realised



depends on the attitude of the observer. No theory explaining understanding by means of experience alone or of reasoning by analogy, could be correct; the impression received being clearly located in the perceived face, and carrying a meaning even when it was not possible to the observer to state exactly *what* meaning. Reflectory imitation, too, being very likely no imitation of movement but of expressive content, can never be explained by those older theories. Some more intrinsic relation between expression and impression must be sought.

As for the apprehensional process proper, we met the paramount importance of apprehension of tension, a fact which is in contradiction to a passive and purely receptive theory of apprehension, since it probably makes for a participation of the observer in perceiving the expressions.

#### ZUSAMMENFASSUNG

Der Zweck dieser Untersuchung ist das Studium einiger Probleme, die sich bei der Deutung des Gesichtsausdrucks der Gemütsbewegungen ergeben. In erster Linie wird der Einfluss der dynamischen Seite des Gesichtsausdrucks untersucht.

Zwei Personen wurden während, hauptsächlich absichtlich hervorgerufenen, emotionellen Situationen gefilmt. Es wurden 65 Filmaufnahmen gemacht, die je einige Sekunden dauerten. Aus jedem wurde der Höhepunkt des Ausdrucks gewählt und als isolierte Photographie kopiert. Die Photographien und Filme wurden 40 Versuchspersonen vorgeführt, welche diese zu deuten hatten. Diese Deutungen wurden nach einer 5-Punkte-Skala beurteilt, nach ihrer Übereinstimmung mit der Introspektion der gefilmten Personen.

Die Resultate zeigen erstens, dass eine korrekte Deutung des Gesichtsausdrucks ohne Zweifel möglich ist; zweitens, dass der Erfolg bei Filmaufnahmen grösser ist als bei statischen Photographien. Die Summe der anerkannten Punkte — ausgedrückt im Prozentsatz des Maximums — beträgt 44.8 % für die Filme und 30.6 % für die Photographien.

Der Charakter der Antworten zeigt, dass die objektive Situation, in welcher sich die wahrgenommene Person befindet, niemals die eigentliche Ursache der Deutung ist; sie hat jedoch einen wesentlichen Einfluss auf das Zustandekommen der Deutung, insbesondere durch die Ausbildung einer Strukturierung



der Ausdrucksergebnisse und durch die Präzisierung des Gefühls, welches dem Ausdruck zugrunde liegt.

Es wurde klar, dass die Bedeutung der Gestalt eines Ausdrucks, d.h. die Konfiguration des Gesichts im gegebenen Moment (wie er sich z.B. auf einer Photographie zeigt), nicht als Ausdruck eines Gefühls gewertet werden kann. Sie ist eben der Ausdruck einer Bereitschaft zur Handlung, oder einer spezifischen Beziehung zur Umgebung oder zum Erlebnis — einer „Positionalität“. Der spezifische Inhalt dieser Beziehung (z.B. eine Bemerkung, eine schmerzliche Erinnerung) ist ohne Bedeutung für den Ausdruck, obwohl er an der Formung des Erlebnisses wesentlichen Anteil hat.

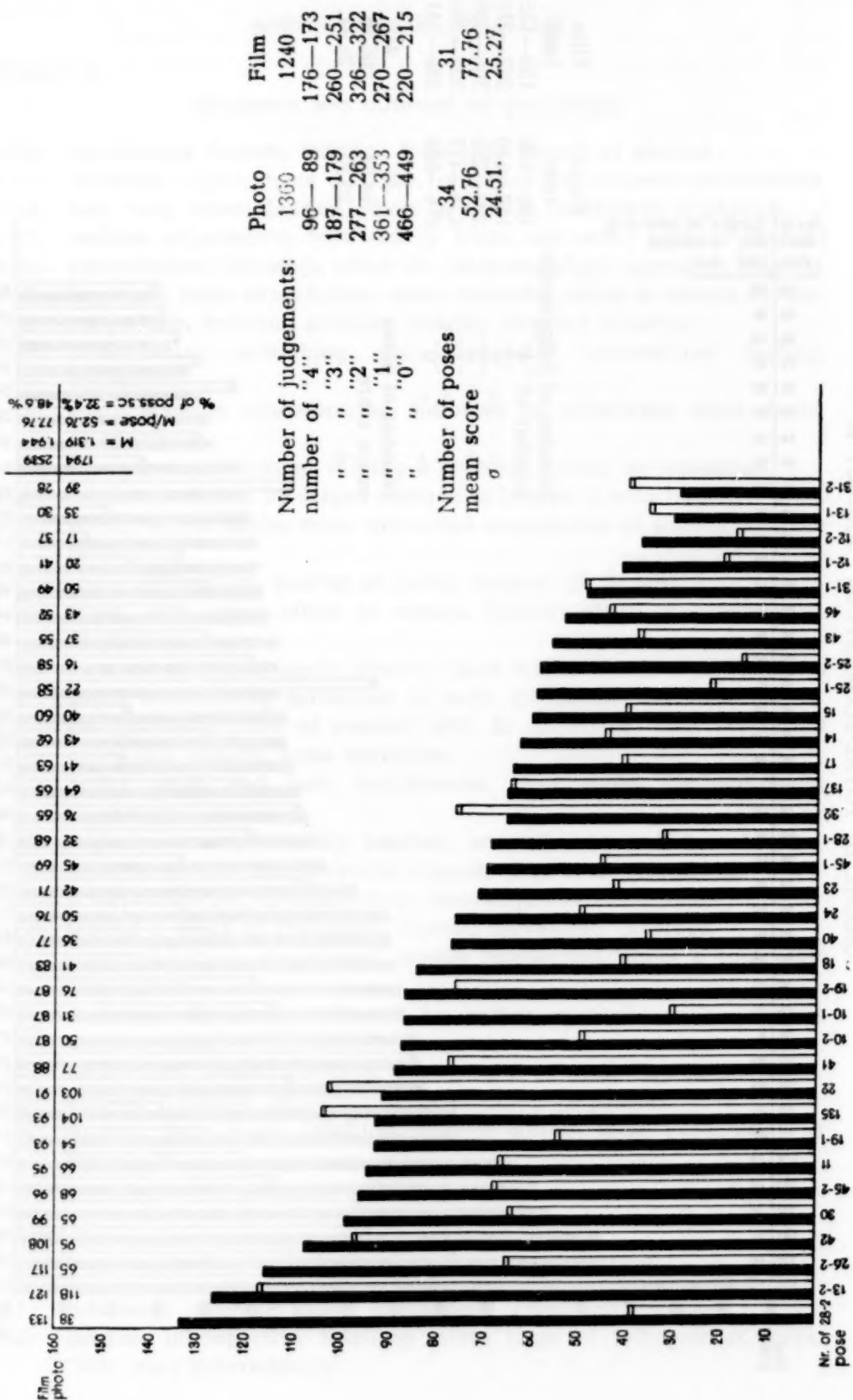
Die dynamische Seite des Ausdrucks, sein Verlauf, ist das expressive Korrelat der Art und Tiefe der Bedeutung, die das Gefühl für den Person hat.

Die Analyse der fehlerhaften Deutungen zeigt, dass diese nicht durch eine wirkliche Mehrdeutigkeit des Ausdrucks verursacht sind. Sie sind vielmehr die Folge der relativen Allgemeinheit der bezüglichen Positionalität (welche z.B. das Erstaunen mit dem Schrecken auf der Photo verwechseln lässt), oder dann ist es die Folge des Umstandes, dass der Deuter ein oder mehrere weniger ausgeprägte Züge des Ausdrucks ausser Acht lässt. Alle Fehldeutungen können erklärt werden wenn wir sie als Deutung eines Bildes betrachten, das nur denselben allgemeinen Charakter besitzt wie das wirklich gebotene Bild, bei dem jedoch einige feinere Einzelheiten fehlen. Diese Einzelheiten scheinen vorwiegend diejenige zu sein, welche die Spannung des Gesichtes anzeigen.

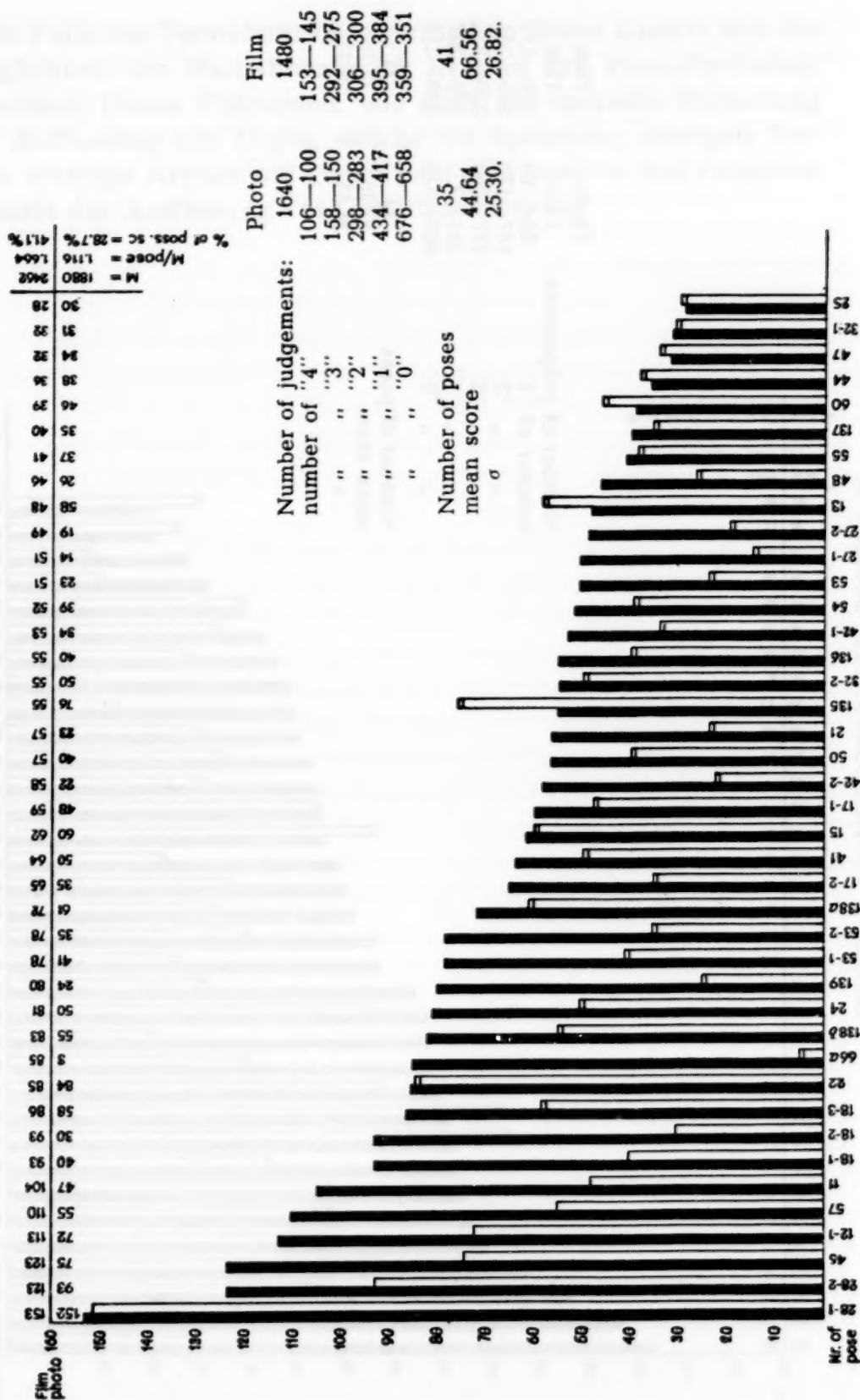
Die Introspektionen der Versuchspersonen erlauben die Schlussfolgerung, dass der Wahrnehmer einen unmittelbaren Eindruck bekommt, welcher er als vom Bilde herausgehend erlebt. Dieser Eindruck wird aber erst in zweiter Instanz expliziert und formuliert. Der Wahrnehmer empfängt jedoch nicht immer einen unmittelbaren Eindruck; die Reaktionen auf einen Ausdruck können verschiedenartig sein. Es können vier Formen unterschieden werden: 1. Die unwillkürliche Nachahmung; 2. die Gefühlsansteckung; 3. die eigene Reaktion des Wahrnehmers auf einen physiognomischen Charakter; 4. Das Verstehen eines bewusst erlebten Eindrucks. Welche dieser Möglichkeiten realisiert wird, ist von der Einstellung des Wahrnehmers abhängig.

Im Falle des Verstehens im eigentlichen Sinne äussert sich die Möglichkeit der Nachahmung oft in eine Art kinaesthetischer Resonanz. Dieses Phänomen, wie auch die spezielle Bedeutung der Auffassung von Zügen, welche die Spannung anzeigen, formen wichtige Argumente gegen jede rein passive und rezeptive Theorie der Auffassung des Gemütsausdrucks.

GRAPH. I: TOTAL-SCORES/POSE SUBJECT A.



GRAPH. II: TOTAL-SCORES/POSE SUBJECT B.





## APPENDIX: LIST OF FILMS

### *Subject A.*

(sequence was different on projection)

- A-10: questioning disgust; tension; somewhat forced. (2 photos)
- A-11: moderate fright; some surprise, preceded by amused conversation.
- A-12: fear; very tense; distrust; effort to remain indifferent. (2 photos)
- A-13: anxious expectation, followed by fright and some pain. (2 photos.)
- A-14: concentrated listening; effort for understanding; approval; interest.
- A-15: impatient tense expectation; some curiosity; effort to remain distant.
- A-17: very tense; resisting attitude; sharply directed attention.
- A-18: remembering something sad; somewhat externalized feeling; thoughtful.
- A-19: sharp acoustic concentration; followed by mimicking slight doubt. (2 photos)
- A-20: concentrated thinking, solving a problem, trying to remember.
- A-23: interest, followed by forced attitude of interest (photo from last part).
- A-24: polite, dutiful smile; some interested expectation of something more interesting.
- A-25: some indignation; feeling of being cheated. (2 photos)
- A-26: tense, with some effort to remain distant; show of indifference. (2 photos)
- A-28: amused attention, some tension, then slight alarm, light anxiety; slight withdrawing movement of body. (2 photos)
- A-30: thoughtful; attitude of contact with E; interested, then thoughtful deliberate speech; some inhibition.
- A-31: polite smile and nod; indifference, then slight disappointment (2 photos).
- A-32: slight melancholy; slightly touched; amused; inhibited, forced.
- A-135: some alertness; then physical disgust; then some indignation.
- A-137: show of sorrow, on instruction. Passivity, forced attitude.
- A-140: tasting a piece of chocolate. Tasting movement; pleasure.
- A-41: some inhibition; amusement, mostly intellectual, about a somewhat risqué joke.
- A-42: depressed listening, followed by rather spontaneous amusement about a joke.
- A-43: melancholy, intent listening, background of some depression.
- A-45: amused listening; appreciation of mockery (2 photos).
- A-46: serious, dark mood; thinking, passive sitting; some expectation of something.

### *Subject B.*

(in almost all poses showing some reticence)

- B-11: interested listening; some distraction; approving reaction.
- B-12: intense, introspective thinking; some tinge of unhappiness; says "No" very determinedly.

- B-13: introverted; remote kind of contact; warm, pleasant and deep feeling.
- B-15: passive thinking; then some surprise, reticent amusement.
- B-17: anxious expectation; withdrawn into herself; trapped feeling; then, dislike, rejection (2 photos).
- B-18: unpleasant sensation, dislike; opposition; decisive rejection. This becomes stronger; effort to remain distant, ending in real dislike and anger (3 photos).
- B-21: fright, preceded by thinking.
- B-22: drawing; concentration on work; some difficulty to overcome.
- B-24: conviction; smiling; then introverted reflection, somewhat depressed.
- B-25: distracted, somewhat depressed.
- B-27: distrust; withdrawal, distance; then forced show of surprised appreciation. Insecurity, some pleasure at end (2 photos).
- B-28: untained amusement, spontaneous laughter; at the end, inhibited. (2 photos)
- B-32: listening to a poem; real gladness; attention; at the end, warm contact-smile towards E (2 photos).
- B-135: concentrated work, then puzzled look at E (photo of puzzledness).
- B-136: looking with interest, some feeling of inferiority.
- B-137: grimace, on instruction; some amusement.
- B-138a: distantiation, expectant, some tension; some timidity-amusement; nervous, involved.
- B-138b: embarrassed; surprised; amused; distant; involved; timidity-laughter.
- B-139: slight amusement; thoughtful, embarrassed, feeling fooled; then shedding sensations.
- B-41: amusement for a corny joke; relaxed; some real amusement.
- B-42: distrust; resistent attitude; expectant; withdrawal; becomes more involved (2 photos).
- B-43: very tense; strong withdrawal; feeling of being trapped.
- B-44: curiosity, some distrust; beginning of withdrawal.
- B-45: olfactory attention. Then physical disgust.
- B-47: "creative moment"; deep pleasant feeling; receptive concentration.
- B-48: relaxation; pondering; pleasant undefined feeling.
- B-50: slight dislike of something she tastes.
- B-53: suspicion; insecurity; pleased suprise; reticent (2 photos).
- B-54: kindly smiling; sympathy; somewhat forced and inhibited.
- B-55: steady, even physical effort; some distance.
- B-57: sudden pull at a string; sudden physical effort.
- B-59: grimace, on instruction; some reticence.
- B-60: aurical concentration, to hear low whispering; some slight amusement.
- B-66: very tense; involved; feeling of catastrophe; tries to look aloof.
- B-66a: movement of affirmation; pleasant mood; superficial.
- B-51: alarm (speed of film 2/3 real speed).

# BIBLIOGRAPHY \*)

- Allport, G. 1937. *Personality, A Psychological Interpretation*. Holt, New-York.
- Binet, A. et Courtier, B. 1896. *La Vie Emotionnelle*. Anné Ps.
- Boring, E. G. et al. 1939. *Introduction to Psychology*. Wiley, New York.
- Bühler, Ch. und Hetzer, H. 1928. *Das Erste Verständnis für Ausdruck*. Z. Ps. 107.
- Buytendijk, F. J. J. 1948. *Algemene Theorie der Menselijke Houding en Beweging*. Spectrum, Utrecht.
- Buzby, D. E. 1924. *The Interpretation of Facial Expression*. Am. J. Ps. 35.
- Coleman, J. 1949. *Facial Expression of Emotions*. Ps. Mon. 63.
- Darwin, Ch. 1872. *The Expression of Emotions in Man and Animals*. London.
- Dumas, G. 1947. *La Vie Affective*, Alcan, Paris.
- Dusenbergh, H. and Knowler, J. 1938. *Specificity of meaning of Facial Expression*. Quart. J. of Speech 24 (Ps. Abstr.).
- Duyker, H. C. J. 1946. *Extra-linguale Elementen in de Spraak*. N. H. Uitg. Mij, Amsterdam.
- Feleky, A. 1914. *The Expression of Emotions*. Ps. Rev. 21.
- Fernberger, S. W. 1928. *False Suggestion and the Piderit Model*. Am. J. Ps. 40.
- Flach, A. 1928. *Die Psychologie der Ausdrucksbewegung*. Arch. Ges. Ps. 65.
- Frois-Wittmann, J. 1930. *The Judgement of Facial Expression*. J. Exp. Ps. 13.
- Goldstein, K. 1947. *The Organism*. World Book Cy, New York.
- Gesell, A. 1928. *Infancy and Human Growth*. Macmillan, New York.
- Guernsey, M. 1928. *Eine Genetische Studie über Nachahmung*. Z. Ps. 107.
- Hanawalt, N. G. 1944. *Role of the Upper Part and Lower Part in the Expression of Emotions*, J. Gen. Ps. 27.
- Hulin and Katz, D. 1935. *The Frois-Wittmann-pictures*, J. Abn. Soc. Ps. 8.
- Jarden, W. and Fernberger, S. W. 1926. *The Effect of Suggestion on the Judgement of Facial Expression of Emotion*. Am. J. Ps. 37.
- Jaspers, K. 1948. *Allgemeine Psychopathologie*. Springer, Berlin (V).
- Kafka, G. 1937. *Grundsätzliches zur Ausdruckspsychologie*. Acta Ps. 3.
- Katz, D. 1951. *Gestaltpsychology*. Methuen, London.
- Kanner, L. 1933. *Judging Emotions from Facial Expression*. Ps. Mon. 41.
- Klages, L. 1935. *Grundlegung der Wissenschaft vom Ausdruck*. Barth, Leipzig (VI, 1942).
- Kline, L. W. and Johannsen, O. E. 1935. *The Comparative Role of the Face and Face-body-hands as Aids in Identifying Emotions*. J. Abn. Soc. Ps. 29.
- Koffka, K. 1928. *The Growth of the Mind*. Kegan Paul, London.
- , 1935. *Principles of Gestalt Psychology*. Harcourt Brace, New York.
- Köhler, W. 1919. *Zur Psychologie der Schimpanzen*. Ps. Fo. 1.
- , 1929. *Gestaltpsychology*. Liveright, New York. (V11, 1945).
- Landis, C. 1924. *A preliminary Study of Facial Expression*. J. Exp. Ps. 7.
- , 1924. *General Behavior as Facial Expression*. J. Comp. Ps. 4.
- , 1929. *The Interpretation of Facial Expression in Emotion*. J. Genet. Ps. 2.

---

\*) Quotations are from the edition, mentioned in brackets.

- Lersch, P. 1932. *Gesicht und Seele*, Reinhardt, München.
- Lhermitte, J. 1939. *L'Image de notre Corps*. N. Rev. Critique, Paris.
- Munn, N. 1939. Knowledge and Judgement in the Interpretation of facial Expression. *Ps. Bull.* 36.
- Oldfield, R. C. 1941. *Psychology of the Interview*. Methuen, London. (III, 1947).
- Plessner, H. 1941. *Lachen und Weinen*. Francke, Bern.
- Révész, G. 1938. *Die Formenwelt des Tastsinnes I*, Nijhoff, Den Haag.
- Rorschach, H. 1921. *Psychodiagnostik*. Huber, Bern.
- Ruckmick, C. A. 1921. A Preliminary Study of the Emotions. *Ps. Mon.* 30.
- Sartre, J. P. 1935. *Esquisse d'une Theorie des Emotions*. Hermann, Paris.
- Scheler, M. 1923. *Wesen und Formen der Sympathie*. Bonn (V, Frankfurt, 1948).
- Sherman, M. 1927. The Differentiation of Emotional Responses in Infants. *J. Comp. Ps.* 7.
- Smith, M. 1946. *A Simplified Guide to Statistics*. Rinehart, New York.
- Thompson, J. 1941. Development of Facial Expression of Emotion in Blind and Seeing Children. *Arch Ps.* 264.
- Turham, M. 1941. (An Experiment concerning the Interpretation of Facial Expressions) *Istanbul Univ. Yeyinlar* 149 (*Ps. Abstr.*).
- Weizsäcker, V. von 1939. *Der Gestaltkreis*. Thieme, Stuttgart (IV, 1950).
- Woodworth, R. S. 1938. *Experimental Psychology*. Holt, New York.



